

**THESE DE DOCTORAT DE  
L'UNIVERSITE PIERRE ET MARIE CURIE**

Spécialité Chimie Analytique  
Ecole Doctorale Chimie Moléculaire

Présentée par  
**Séverine CLAVIER**

Pour obtenir le grade de  
**DOCTEUR de L'UNIVERSITE PIERRE ET MARIE CURIE**

**Effet moléculaire du peptide vecteur (R/W)<sub>9</sub> sur le phénotype de  
cellules modèles du sarcome d'Ewing : Etude protéomique**

Soutenance le 16 octobre 2014

Devant le jury composé de :

**Dr. Joëlle VINH**, SMBP, *ESPCI*

**Pr. Guillaume VAN DER REST**, LCP, *Université Paris Sud*

**Pr. Tula SAISON-BEHMOARAS**, *Régulation et dynamique des génomes, MNHN*

**Pr. Richard COLE**, IPCM-LCSOB, *UPMC*

**Dr. Jacques GHYSDAEL**, *Signalisation, Neurobiologie et Cancer, Institut Curie*

**Dr. Emmanuelle SACHON**, LBM, *UPMC*

**Dr. Gérard BOLBACH**, LBM, *UPMC*

**Dr. Sandrine SAGAN**, LBM, *UPMC*

Rapportrice

Rapporteur

Examinatrice

Examineur

Examineur

Directrice de thèse

Directeur de thèse

Invitée



## Remerciements

Pour commencer je souhaite vivement remercier Joëlle Vinh et Guillaume van der Rest qui ont accepté d'être les rapporteurs de ce travail de thèse, ainsi que Richard Cole, Tula Saison-Behmoaras et Jacques Ghysdaël pour leur participation à mon jury de soutenance en tant qu'examinateurs.

Merci à Laurence Sabatier, qui à Strasbourg lors des enseignements de la filière chimie analytique puis lors de mon stage M2, m'a fait découvrir le passionnant domaine de la bioanalyse.

Au cours de ces trois belles années, nombreuses sont les personnes avec lesquelles j'ai eu la chance de travailler au sein de la fac de Jussieu et que je souhaite chaleureusement remercier.

Tout d'abord, Manu et Gérard, mes directeurs de thèse, un co-encadrement qui a idéalement fonctionné avec des avis, des conseils toujours très complémentaires. Un grand merci à Gérard de m'avoir, dès qu'une occasion se présentait, fait bénéficier de son expérience en spectrométrie de masse. J'ai beaucoup apprécié nos nombreuses discussions scientifiques ou non qui me permettait de prendre du recul, de faire évoluer mon point de vue pour la réalisation d'expériences et aussi pour des choix de vie. Merci infiniment à Manu, ma complice au cours de ces trois années. C'était formidable de travailler avec toi et ce n'ai pas sans un pincement au cœur que je m'apprête à tourner cette page. Tu t'es toujours montrée très disponible et en toutes circonstances prête à m'aider à faire face aux difficultés rencontrées dans la réalisation de mon projet de thèse. Merci pour ton dynamisme et ton optimisme à toute épreuve ! Je suis sûre que cette amitié forte que nous avons développée au cours de ces trois belles années persistera très très longtemps.

Merci Sandrine pour ton accueil au sein du LBM2, tes précieux conseils scientifiques et ton implication régulière dans mon projet de thèse et cela malgré tes nombreuses responsabilités. Merci pour l'énergie que tu dépenses au quotidien pour maintenir des bonnes conditions de travail et la très agréable ambiance du laboratoire.

Merci Gilles pour ton incroyable gentillesse et bienveillance ! C'était super de t'avoir sur la plateforme. Tu es une personne sur laquelle on peut compter et je te souhaite le meilleur pour le futur.

Merci Claire, Dominique, Karine et Nicolas pour votre accueil toujours chaleureux dans la pièce qui fut jusqu'à quelques jours de cela votre bureau au Bât F. Pendant les migrations des gels ou les incubations sur billes, etc... je ne compte pas le nombre de fois où je suis venue « squatter » chez vous et j'ai apprécié que ce soit autant d'occasions de discuter et d'échanger sur nos projets de recherche ou sur le métier d'enseignant-chercheur. Merci Sophie pour les coups de pouce, les infos utiles...et le chocolat ;-). Merci Jésus pour ta disponibilité et les discussions scientifiques.

Merci à l'équipe de mon deuxième bureau d'accueil lors de mes manips au bât F et du café de l'aprem. A Fabienne pour ses conseils d'experte en internalisation de CPPs et à Françoise pour toutes les techniques de biochimie qu'elle m'a transmise sans oublier l'apprentissage de la culture cellulaire !!

## Remerciements

Merci à la super team de filles, ex-doctorantes LMB2, avec lesquelles j'ai eu la chance de débiter ma thèse : Lucrèce, Astrid, Chérine et Dimitra. C'était génial de partager des verres avec vous à l'Inévitable et de pouvoir continuer à échanger avec vous par mail après votre départ du labo (Est-ce que vous auriez accès à la publi X par hasard ?? ;-)). Mention spéciale pour Lucrèce qui m'a, en quelques sortes, ouvert la porte du LBM (Viens donc faire un tour chez nous si tu veux faire de la bioanalyse !) et qui a continué à m'offrir son aide et ses conseils depuis les Pays-Bas. Je suis contente que tu sois parvenue à revenir à Jussieu !

Un grand merci aussi à la team de gars ayant succédé à cette team de filles : Tom, Pierre et Mehdi. C'était vraiment chouette de se croiser au bât F ou de partager des moments de manip ! Merci Tom d'avoir toujours veillé à me prévenir, lorsque, « exilée » au bât A, je risquais de manquer le départ vers la cantine, les bords de seine ou la rue Linnée. Et merci aux thésards LBM3, Shahid (« Beers tonight Dude ? »), les 2 Anaïs, Lucie, Lucille ainsi qu'aux doctorants et postdoctorants Bessem, Jean, Adrian, Bayie, Quentin de l'IPCM-CSOB voisin pour les bons moments partagés !

Merci à nos voisins du LBM3 et enseignants LC351, Olivier, Ludovic, Emeric et Lucie.

J'ai énormément apprécié l'ambiance du 7<sup>ème</sup> étage : Même avec une déco disons...datée on peut travailler dans la bonne humeur ! Vive les productions cinématographiques du LBM2 et en particulier la « Danse de la salle de culture ».

Merci Rodrigue pour la synthèse et la purification de ce fameux peptide (R/W)<sub>9</sub> sous ses différentes versions, photoactivable, deutérée, de configuration D... C'était parfait et sans cela je ne serais pas allée bien loin !

Merci à Eliane, Nathalie, Stéphane et maintenant Karine pour la gestion administrative du LBM. J'ai apprécié votre disponibilité et votre efficacité.

Merci à tous mes collègues d'enseignements et en particulier à ceux du module LC351 pour les coups de mains, les échanges de créneaux de cours. Merci à Laurent Gaillon pour la formation aux TPs et TDs d'HPLC et à Richard Cole pour ses explications sur certains exos de spectrométrie de masse. Merci encore une fois à Manu qui m'a également fourni une aide précieuse pour la préparation des enseignements !

Côté Bât A, je souhaite remercier Christophe, Tahar ainsi que les membres de l'équipe BIOSIPE. Merci « tonton » Thierry d'avoir été mon tuteur de thèse ! Merci Ali pour ta confiance lors des ateliers du NV601. Merci à Sonia et Mouna pour leur sympathie, les ateliers microbalance et speedvac auraient été nettement moins agréables sans vous !

Merci aussi aux membres de l'IPCM, Jean-Philippe Goddard et Cyril Ollivier ainsi qu'à leurs étudiants qui m'ont permis d'utiliser le système Rayonet.

Au-delà des personnes rencontrées à Jussieu, je souhaite remercier mes collaborateurs et soutiens extérieurs.

Xiuxia Du et Adam Baxter de l'Université de Charlotte en Caroline du Nord qui nous ont d'abord aidés pour les données de cross-linking chimique et qui nous ont ensuite suivi pour la mise au point d'un logiciel pour le photocross-linking. Merci pour leur patience dans le débogage du logiciel et pour les sympathiques discussions Skype.



## Remerciements

Un immense merci à Jacques Ghysdaël de l'Institut Curie. Sans toi, pas de cellules EF! Merci pour tous les dépannages et les conseils. Je te suis extrêmement reconnaissante pour l'aide que tu m'as apportée tout au long de ma thèse. Merci Jacques !

Merci beaucoup aux bioinformaticiens de l'Institut Curie, Patrick Pouillet et Guillaume Arras pour m'avoir permis d'utiliser MyProMS. Merci Patrick d'avoir répondu à mes innombrables emails et d'avoir toujours été en mesure de régler les problèmes informatiques.

Merci à la chouette équipe du SMBP de l'ESPCI et en particulier à mes amis Shaker et Sophie pour leurs dépannages Nano ! (Un bout de capillaire de 20µm à me prêter ? une idée pour le passeur d'échantillons ??)

Merci à l'Ecole Doctorale Iviv qui m'a attribuée mon financement de thèse. Merci en particulier à Germain pour son investissement dans la vie de l'ED avec les journées de rentrée et la super formation BioNoBio. Malheureusement, les belles choses ne durent pas toujours...

Je souhaite également remercier la Société Française de Spectrométrie de Masse (SFSM) qui a financé ma participation à deux congrès (à Orléans en 2012 et en Suisse en 2014) au cours de ma thèse et qui m'a permis de m'engager dans une belle aventure associative via le bureau du CJSM. Merci Amandine de m'avoir donné l'envie de m'investir dans le bureau et de m'avoir montrée la voie et merci aux membres actifs de l'association faisant partie du bureau ou non pour les supers moments partagés lors des congrès ou des RCJSMs !!

Merci à Martine, ma prof de danse du 5<sup>ième</sup> et à ses danseuses. Ces cours géniaux vont tellement me manquer...

Finalement (last BUT not least...) je voudrais remercier tous ceux qui ont suivi et participé à cette aventure parisienne : mes fantastiques amis de lycées, Déborah, Maé, Kana, Steven, Jacqueline, Ramzi, Pauline et les supers pièces rapportées ;-) Sylvain et Matthieu, mais aussi Goulven et Kristen, Emilie et Loic, et puis aussi les supers de l'ECPM, Camille, Caro, Shaker, Sophie, Pauline, Manon (et leurs +1). Sans oublier Nadège, Thomas, Mathilde et Yoann. Si j'ai passé de si belles années c'est aussi grâce à vous !!

Merci infiniment à mes parents qui m'ont toujours soutenue, réconfortée, conseillée... déménagée ! Merci pour votre amour et merci à ma soeurette et mamie adorées ! Merci aussi à ma géniale belle famille ☺

Enfin, merci à toi, mon petit Lou qui rend la vie si belle...



# Liste des publications et communications

## Publications :

**Clavier S**, Du X, Sagan S, Bolbach G, Sachon E. « *An integrated cross-linking-MS approach to investigate cell penetrating peptides interacting partners.* » EuPA Open Proteomics, Vol 3, 2014 June

Matheron L, **Clavier S**, Diebete O, Karoyan P, Bolbach G, Guianvarc'h D, Sachon E « *Improving the selectivity of the phosphoric acid B-elimination on a biotinylated phosphopeptide* » J Am Soc Mass Spectrom. 2012 Nov 23(11):1981-90

Christine C, Koubemba M, Shakir S, **Clavier S**, Ehret-Sabatier L, Saupe F, Orend G, Charbonnière LJ. *"Synthesis of an activated phosphonated bifunctional chelate with potential for PET imaging and radiotherapy."* Org Biomol Chem. 2012 Dec 14;10(46):9183-90

## Communications :

### Orales :

- *"Understanding the molecular basis of (R/W)<sub>9</sub> CPP effect on tumoral EF cells phenotype "* (SMAP 2014, Lyon, Congrès national).
- *"Etudes des partenaires intracellulaires du peptide vecteur (R/W)<sub>9</sub> par pontage covalent et spectrométrie de masse"* (Rencontres du Club des Jeunes Spectrométristes de Masse, Pau, RCJSM 2013).
- *"Chemical Cross-linking and Mass Spectrometry for the study of (R/W)<sub>9</sub> Cell Penetrating Peptide interactions with its partner proteins"* (JFSM 2012, Orléans, Congrès national)

### Posters :

- *"Molecular effect of (R/W)<sub>9</sub> Cell Penetrating Peptide on Ewing's sarcoma model cells phenotype : Proteomic study"* (IMSC 2014, Genève, Congrès international). **Séverine Clavier**, Jacques Ghysdaël, Sandrine Sagan, Gérard Bolbach and Emmanuelle Sachon
- *"Understanding (R/W)<sub>9</sub> Cell Penetrating Peptide effect on tumoral cells: Association of targeted and differential approaches "* (Eupa 2013, St Malo, Congrès européen). **Séverine Clavier**, Rodrigue Marquant, Xiuxia Du, Sandrine Sagan, Gérard Bolbach and Emmanuelle Sachon.

## *Liste des publications et communications*

## Liste des abbréviations et acronymes

ABP	actin binding protein
ACN	Acétonitrile
ADN/ARN	acide déoxyribonucléique/ acide ribonucléique
AP-MS	affinity purification-mass spectrometry
ATP/ADP	Adenosine triphosphate/ Adenosine diphosphate
BCA	BiCinchoninic acid Assay
BN-PAGE	Blue native polyacrylamide gel electrophoresis
BSA	Bovin serum albumin
CHCA	$\alpha$ -Cyano-4-hydroxycinnamic acid
CHO	Chinese Hamster Ovary
CID	collision induced dissociation
CPP	Cell Penetrating Peptide ou peptide vecteur
DAPI	4',6'-diamidino-2-phénylindole
DAVID	Database for Annotation Visualization and Integrated Discovery
DDA	Data Dependent Acquisition
DIA	Data Independent Acquisition
DMEM	Dulbecco's Modified Eagle's Medium
DTT	Dithiothréitol
ECD	Electron capture dissociation
ESI	electrospray ionisation
ETD	electron transfert dissociation
FA	formic acid
FASP	Filter aided sample preparation
FBS	Fœtal bovin serum
FDR	false discovery rate
FITC	fluorescein isothiocyanate
FT-ICR	fourier transform-ion cyclotron resonance
GAG	Glycosaminoglycanes
GO	Gene ontology
GTP	Guanosine triphosphate
HBSS	Hank's Balanced Salt Solution
HCl	Hydrochloric acid
HDX	hydrogen-deuterium exchange
HPLC	High Performance Liquid chromatography
HRP	Horseradish peroxidase
IAA	Iodoacetamide
IEF	isoelectrofocalisation
ITC	isothermal titration calorimetry
LTQ	Linear trap quadrupole
MALDI	matrix assisted laser desorption/ionisation
MS	mass spectrometry
NCS	Newborn calf serum
NSD	nozzle skimmer dissociation

## *Liste des abréviations et acronymes*

OFP	oxidative footprinting
PBS	Phosphate buffered saline
PSMs	Peptide-to-spectrum matches
PTM	Post-translational modification
RMN	résonance magnétique nucléaire
ROS	Reactive oxygen species
SAINT	Significance Analysis of Interactome
SDS-PAGE	Sodium dodecyl sulfate- polyacrylamide gel electrophoresis
SILAC	Stable Isotope Labelling with Amino acids in Cell culture
SPE	Solid phase extraction
SPR	surface plasmon resonance
STRING	Search Tool for the Retrieval of Interacting Genes/Proteins
SWATH	Sequential Window Acquisition of all Theoretical spectra
TFA	trifluoroacetic acid
TOF	Time of Flight
UV	ultraviolet

## Tables des matières

<b>Remerciements</b>	- 3 -
<b>Liste des publications et communications</b>	- 7 -
<b>Liste des abréviations et acronymes</b>	- 9 -
<b>Liste des figures et tableaux</b>	- 15 -
<b>Préface</b>	- 21 -
<b>Chapitre 1 : Des interactions protéine-protéine au phénotype cellulaire, l'exemple du cytosquelette d'actine</b>	- 23 -
<b>I. Au sujet des interactions protéine-protéine</b>	- 24 -
1. Les différents types d'interactions protéine-protéine	- 24 -
2. Les domaines d'interactions des protéines	- 24 -
3. Régulation et effets biologiques des interactions protéine-protéine	- 25 -
4. Les réseaux d'interactions des protéines	- 26 -
<b>II. Le cytosquelette d'actine : illustration de la complexité des réseaux d'interactions</b>	- 27 -
1. L'actine : fiche d'identité [19]	- 27 -
2. Polymérisation de l'actine	- 28 -
3. Les protéines de liaison de l'actine et leurs rôles	- 29 -
4. Le rôle du cytosquelette dans la réponse de la cellule aux contraintes mécaniques externes	- 33 -
<b>III. Cytosquelette et oncogénèse</b>	- 34 -
1. Généralités [72]	- 34 -
2. Sarcome d'Ewing : Comment la protéine EWS-FLI1 agit-elle sur le cytosquelette ?	- 35 -
<b>Chapitre 2 : Méthodes pour l'étude des interactions protéine-protéine : de la découverte de nouveaux partenaires à l'intégration dans les réseaux d'interactions</b>	- 39 -
<b>I. Méthodes ciblées de découverte de partenaires d'interaction</b>	- 41 -
1. Systèmes cellulaires d'expression de double hybride	- 41 -
2. Résonance plasmonique de surface ( <i>Surface plasmon resonance</i> , SPR)	- 42 -
3. Blue Native PAGE	- 45 -
4. Méthodes de purification d'affinité couplées à la spectrométrie de masse (AP-MS)	- 45 -
5. Pontage covalent ou "Cross-linking" couplé à la spectrométrie de masse	- 55 -
<b>II. Méthodes de validation et de caractérisation</b>	- 59 -
1. Calorimétrie de titrage isotherme (ITC)	- 59 -
2. Résonance magnétique nucléaire (RMN)	- 61 -
3. Cristallographie aux rayons X	- 63 -
4. Spectrométrie de masse native	- 63 -
5. L'échange hydrogène/deutérium (HDX)	- 65 -
6. Le marquage par oxydation (OFP)	- 67 -
<b>III. Approches de protéomique différentielle quantitative pour l'étude des réseaux d'interactions</b>	- 68 -
1. Présentation des principales méthodes de quantification en protéomique différentielle	- 68 -
2. Intérêt pour les études d'interactions protéine-protéine	- 73 -
<b>IV. Complémentarité des différents outils d'étude des interactions protéine-protéine</b>	- 74 -
<b>V. Perspectives : Vers une étude dynamique possible des interactions protéine-protéine, exemple de la technique AP-SWATH</b>	- 76 -
<b>Chapitre 3 : Les peptides vecteurs</b>	- 79 -
<b>I. Présentation générale des peptides vecteurs :</b>	- 79 -

1.	Découverte des CPPs	- 79 -
2.	Les caractéristiques des CPPs et leur classification	- 79 -
<b>II.</b>	<b>Mécanisme d'internalisation et outils d'études</b>	<b>- 81 -</b>
1.	Description des mécanismes proposés pour l'internalisation des CPPs	- 81 -
2.	Méthodes d'études de la capacité d'internalisation des CPPs dans les cellules	- 86 -
<b>III.</b>	<b>Seulement des peptides vecteurs ?</b>	<b>- 88 -</b>
1.	Les domaines d'applications des CPPs	- 88 -
2.	L'influence des cargos sur le devenir intracellulaire des conjugués	- 88 -
3.	Des peptides vecteurs avec une activité biologique intrinsèque	- 89 -
	<b>Présentation du projet de recherche</b>	<b>- 91 -</b>
<b>I.</b>	<b>Contexte de l'étude</b>	<b>- 91 -</b>
<b>II.</b>	<b>Objectif et structure de ce projet de thèse</b>	<b>- 93 -</b>
	<b>Chapitre 4 : Cross-linking chimique couplé à la spectrométrie de masse pour étudier l'interaction des CPPs (R/W)<sub>9</sub> et (R/W)<sub>16</sub> avec l'actine.</b>	<b>- 95 -</b>
<b>I.</b>	<b>Objectifs de l'article 1</b>	<b>- 95 -</b>
<b>II.</b>	<b>Principaux résultats</b>	<b>- 95 -</b>
1.	Compositions des complexes	- 95 -
2.	Zones d'interaction	- 96 -
	<b>Chapitre 5 : Vers la capture in cellulo de partenaires d'interaction du peptide vecteur (R/W)<sub>9</sub> par photocross-linking et spectrométrie de masse</b>	<b>- 97 -</b>
<b>I.</b>	<b>Mise au point et validation biologique d'une version photoactivable du CPP (R/W)<sub>9</sub> : photo(R/W)<sub>9</sub></b>	<b>- 97 -</b>
1.	Design et synthèse de photo(R/W) <sub>9</sub>	- 97 -
2.	Validation biologique de photo(R/W) <sub>9</sub>	- 99 -
<b>II.</b>	<b>Validation de la faisabilité de la réaction de photocross-linking in vitro et mise au point d'outils</b>	<b>- 106 -</b>
1.	Réaction de photocross-linking in vitro	- 106 -
2.	Analyses des espèces photocross-linkées par spectrométrie de masse MALDI-TOF/TOF et nanoLC-ESI-LTQ-Orbitrap.	- 108 -
<b>III.</b>	<b>Expériences de photocross-linking in cellulo</b>	<b>- 126 -</b>
1.	Evaluation de la cytotoxicité de l'irradiation	- 126 -
2.	Expériences de photocross-linking in cellulo	- 126 -
	<b>Chapitre 6 : Pull down et photocross-linking</b>	<b>- 139 -</b>
<b>I.</b>	<b>Présentation de l'expérience de pull-down associée au photocross-linking</b>	<b>- 139 -</b>
<b>II.</b>	<b>Résultats obtenus</b>	<b>- 141 -</b>
1.	Electrophorèse sur gel 1D-SDS-PAGE	- 141 -
2.	Résultats des analyses de spectrométrie de masse haut débit.	- 141 -
<b>III.</b>	<b>Interprétation des données : Quelles sont les « vraies » interactions?</b>	<b>- 142 -</b>
1.	Génération des listes de partenaires d'interaction potentiels	- 142 -
2.	Interprétation avec les outils d'enrichissement de termes de <i>gene ontology</i>	- 144 -
3.	Examen manuel des listes de partenaires d'interaction potentiels	- 145 -
<b>IV.</b>	<b>Signification biologique des résultats obtenus</b>	<b>- 146 -</b>
1.	Une interaction avec la protéine EWS	- 146 -
2.	Possibilité d'une interaction avec EWS-FLI1	- 147 -



**Chapitre 7 : Evaluation de l'effet de (R/W)<sub>9</sub> sur le protéome des cellules EF par une approche de protéomique quantitative différentielle.** \_\_\_\_\_ - 151 -

- I. Mise en œuvre de la technique SILAC.** \_\_\_\_\_ - 151 -
  - 1. Vérification du taux d'incorporation des acides aminés marqués. \_\_\_\_\_ - 151 -
  - 2. Détermination des seuils de significativité pour la variation des ratios H/L \_\_\_\_\_ - 153 -
- II. Comparaison des protéomes des cellules EF et 3T3.** \_\_\_\_\_ - 155 -
  - 1. Protéines significativement surexprimées dans les cellules EF par rapport aux cellules 3T3. \_\_\_\_ - 157 -
  - 2. Protéines significativement sous-exprimées dans les cellules EF par rapport aux cellules 3T3. \_ - 159 -
- III. Effet du peptide vecteur photo(R/W)<sub>9</sub> sur le protéome des cellules EF.** \_\_\_\_\_ - 162 -

**Conclusion** \_\_\_\_\_ - 167 -

**Matériel et Méthodes** \_\_\_\_\_ - 171 -

- I. Matériel** \_\_\_\_\_ - 171 -
  - 1. Origine du matériel \_\_\_\_\_ - 171 -
  - 2. Composition de tampons et solutions couramment utilisés : \_\_\_\_\_ - 171 -
- II. Techniques d'analyses et culture cellulaire** \_\_\_\_\_ - 172 -
  - 1. Spectrométrie de masse MALDI-TOF \_\_\_\_\_ - 172 -
  - 2. Couplage NanoLC-ESI-MS/MS \_\_\_\_\_ - 172 -
  - 3. Culture cellulaire \_\_\_\_\_ - 173 -
- III. Protocoles** \_\_\_\_\_ - 173 -
  - 1. Evaluation de la cytotoxicité de peptides vecteurs \_\_\_\_\_ - 174 -
  - 2. Evaluation de la quantité de peptide internalisé \_\_\_\_\_ - 175 -
  - 3. Observation du cytosquelette d'actine des cellules par immunofluorescence \_\_\_\_\_ - 176 -
  - 4. Etude *in vitro* de la réaction de photo-crosslinking \_\_\_\_\_ - 176 -
  - 5. Expérience de photocross-linking avec la stratégie N°1 \_\_\_\_\_ - 177 -
  - 6. Réalisation d'un western blot pour détecter les espèces biotinylées \_\_\_\_\_ - 179 -
  - 7. Expérience de photocross-linking avec la stratégie N°2 \_\_\_\_\_ - 179 -
  - 8. Expérience de pull-down + photocross-linking \_\_\_\_\_ - 180 -
  - 9. Expériences de protéomique différentielle avec la technique SILAC (*Stable Isotope Labelling with Amino acids in cell Culture*) \_\_\_\_\_ - 181 -

**Références bibliographiques** \_\_\_\_\_ - 185 -

**Annexes** \_\_\_\_\_ - 205 -

- I. Annexe instrumentale : Spectrométrie de masse** \_\_\_\_\_ - 207 -
  - 1. Spectromètres de masse : Généralités \_\_\_\_\_ - 207 -
  - 2. Représentations schématiques des spectromètres de masse utilisés \_\_\_\_\_ - 212 -
- II. Annexes diverses** \_\_\_\_\_ - 215 -



## Liste des figures et tableaux

Figure 1 : Représentation schématique de la structure cristallographique d'un monomère d'actine.....	28 -
Figure 2 : Schéma de la polymérisation de l'actine.....	28 -
Figure 3 : Représentation schématique de la dynamique de polymérisation de l'actine.....	30 -
Figure 4 : Les trois arrangements de l'actine .....	32 -
Figure 5 : Représentation schématique de l'obtention de la protéine de fusion EWS-FLI1 par translocation chromosomale. ....	36 -
Figure 6 : Frise chronologique récapitulant les évolutions de la recherche sur les interactions protéine-protéine.....	39 -
Figure 7 : Schéma récapitulant les différentes méthodes d'étude d'interactions protéine-protéine qui seront abordées .....	40 -
Figure 8 : Exemple de stratégie utilisée pour détecter une interaction protéine-protéine avec un système double hybride rétablissant l'activité de la protéine GAL4. ....	41 -
Figure 9 : Schéma de principe de la résonance plasmonique de surface (SPR) .....	43 -
Figure 10 : Représentation schématique d'une mesure de SPR .....	44 -
Figure 11 : Description schématique des étapes universelles d'une expérience d'immunoprécipitation couplée à la spectrométrie de masse [113]: .....	46 -
Figure 12 : Obtention d'une protéine recombinante porteuse d'une étiquette .....	48 -
Figure 13 : Schéma de l'expérience de TAP-tag initiale .....	49 -
Figure 14 : Réaction de cross-linking chimique.....	55 -
Figure 15 : Structure du formaldéhyde PFA et du glutaraldéhyde GA.....	56 -
Figure 16 : Structure chimique des acides aminés photoactivables leucine et méthionine et de leurs équivalents naturels.....	58 -
Figure 17 : Configuration classique d'un calorimètre de titrage isotherme .....	60 -
Figure 18: (A) Données brutes extraites d'une expérience d'ITC. (B) Isotherme de liaison représentant l'échange de chaleur en fonction du ratio molaire ligand/protéine.....	60 -
Figure 19 : Exemple de spectres $^1\text{H}$ - $^{15}\text{N}$ HSQC de l'interleukin-2 (IL-2) en présence (bleu) et en l'absence (marron) d'une petite molécule inhibitrice de IL-2.....	62 -
Figure 20 : Spectres de masse obtenus en infusion avec une source électrospray et permettant de suivre l'hydrolyse de l'hexasaccharide de N-acétylglycosamine (NAG <sub>6</sub> ) par le lysozyme de blanc d'œuf de poule (HEWL).....	64 -
Figure 21 : Graphiques donnant le taux de deutération de différents domaines de la protéine apo-myoglobine pour une expérience d'HDX-MS.....	67 -
Figure 22: Schéma des protocoles de spectrométrie de masse quantitative indiquant à quelle étape les deux conditions d'une expérience sont rassemblées et donc à quel niveau l'introduction d'erreurs systématiques dans la quantification liées à la manipulation d'échantillon est arrêtée. Les rectangles bleus et jaunes correspondent aux deux conditions expérimentales à comparer.....	69 -
Figure 23 : Exemple de protocole pouvant être utilisé pour réaliser la quantification relative de protéines par la technique SILAC. ....	70 -
Figure 24 : Image d'un gel 2D réalisée avec la technique DiGE .....	71 -
Figure 25: Schéma du concept de biologie en réseaux d'interactions (Network biology).....	74 -

Figure 26 : Roue d'Edmunson : projection sous forme d'hélice alpha de la séquence (R/W) <sub>9</sub> .....	- 80 -
Figure 27 : Formation de liaisons hydrogènes divalentes entre les groupements guanidinium des chaînes latérales des arginines et des groupements phosphates, sulfates et carboxylates .....	- 82 -
Figure 28 : Les différentes voies d'entrée par endocytose dans la cellule .....	- 84 -
Figure 29 : Modèle proposé pour expliquer la translocation directe des CPPs par formation de micelles inverses.....	- 85 -
Figure 30 : Schéma des différentes étapes de translocation directe par formation de pores. ....	- 86 -
Figure 31 : Analyse quantitative des effets de différents peptides vecteurs sur la morphologie des cellules 3T3 et EF .....	- 92 -
Figure 32: Structure développée du peptide photo(R/W) <sub>9</sub> mettant en évidence les fonctions de ses différentes parties constitutives. ....	- 98 -
Figure 33 : Signal obtenu en spectrométrie de masse MALDI-TOF pour un mélange 1/1 des versions deutérées et non deutérées du peptide photo(R/W) <sub>9</sub> .....	- 98 -
Figure 34 : Mécanisme de transformation du WST-8 (composé légèrement jaune) en WST-8 formazan (composé orange) en présence de cellules viables .....	- 99 -
Figure 35 : Pourcentage de viabilité des cellules EF à l'issue de 2 h, 3h30 et 5 h d'incubation avec le peptide photo(R/W) <sub>9</sub> pour des concentrations comprises entre 1 et 30 µM.....	- 100 -
Figure 36 : Les différentes étapes du protocole de quantification de l'internalisation d'un peptide vecteur (CPP). ....	- 102 -
Figure 37 : Résultats obtenus pour les expériences de quantification de l'internalisation des peptides photo(R/W) <sub>9</sub> et (R/W) <sub>9</sub> réf dans les cellules EF à 37 °C pour des concentrations extracellulaires en peptides de 5 et 7,5 µM et une période d'incubation de 75 min. ....	- 103 -
Figure 38 : Images de microscopie de fluorescence réalisées pour confirmer l'effet du peptide photo(R/W) <sub>9</sub> sur le remodelage du cytosquelette d'actine dans les cellules EF.....	- 105 -
Figure 39 : Représentation schématique du protocole utilisé pour étudier les peptides cross-linkés issus de la réaction de photocross-linking du peptide photo(R/W) <sub>9</sub> avec la BSA. ....	- 106 -
Figure 40 : Mécanisme de la réaction de photocross-linking entre une sonde de type benzophénone et une protéine.....	- 107 -
Figure 41 : Schéma du montage de photoirradiation UV avec une lampe à tungstène. ....	- 107 -
Figure 42 : Spectre MALDI-TOF (gamme de m/z 1000-3000) obtenu pour le milieu réactionnel de photocross-linking entre photo(R/W) <sub>9</sub> et la BSA à l'issue de la digestion trypsique et de la purification d'affinité. ....	- 108 -
Figure 43 : Spectre de fragmentation MALDI-TOF/TOF du doublet de peptides photocross-linkés 2803-2813. ....	- 110 -
Figure 44 : Représentation schématique de la structure potentielle de l'ion du peptide photocross-linké correspondant au doublet 2803-2813.....	- 111 -
Figure 45 : Peptides correspondant à la protéine BSA (P02769) identifiés lors de la recherche avec Mascot de l'analyse NanoLC-ESI-MS/MS du milieu de photocross-linking photo(R/W) <sub>9</sub> -BSA. ....	- 113 -
Figure 46 : Liste de candidats obtenus avec Xlink-Identifier pour le peptide photocross-linké (M+H) <sup>+</sup> m/z 2803.....	- 114 -
Figure 47 : Schéma des différents fragments b et y (fragmentation CID) pouvant être obtenus à partir d'un ion précurseur composé de deux peptides liés covalamment par photocross-linking.. ....	- 115 -
Figure 48 : Exemple de fichier candidat Xlink-Identifier 2.0 répertoriant les pics identifiés dans un spectre MS/MS donné.....	- 116 -

Figure 49 : Procédure d'identification du (ou des) résidu(s) potentiellement impliqué(s) dans la réaction de photocross-linking.....	- 117 -
Figure 50 : Représentation des sites de fragmentation d'un peptide selon la nomenclature de Biemann.....	- 119 -
Figure 51 : Production du fluoranthène radical anion dans le volume d'ion.....	- 119 -
Figure 52: Récapitulatif des ions observés en CID basse énergie (ESI-LTQ/orbitrap) pour l'ion photocross-linké de masse moléculaire 2803-2813 chargé 4+.....	- 121 -
Figure 53 : Récapitulatif des ions observés en ETD (avec énergie supplémentaire de 20 %) (ESI-LTQ/orbitrap) pour l'ion photocross-linké de masse moléculaire 2803-2813 chargé 4 <sup>+</sup> .....	- 122 -
Figure 54 : Récapitulatif des ions observés en CID haute énergie (MALDI-TOF/TOF) pour l'ion photocross-linké de (M+H) <sup>+</sup> 2803-2813 chargé 1 <sup>+</sup> .....	- 123 -
Figure 55 : Stratégie initiale (N°1) envisagée pour l'identification des partenaires d'interaction (intracellulaires et/ou membranaires) du peptide (R/W) <sub>9</sub> par photocross-linking.....	- 127 -
Figure 56: Gel 1D SDS-PAGE coloré au bleu de Coomassie pour l'expérience de photocrosslinking <i>in cellulo</i> avec le peptide photo(R/W) <sub>9</sub> . ....	- 128 -
Figure 57 : Spectre MALDI-TOF obtenu pour l'éluat correspondant au milieu réactionnel v digéré et purifié sur billes.....	- 129 -
Figure 58 : Image de 1) gel 1D SDS-PAGE coloré au bleu de Coomassie et de 2) western-blot réalisé avec de la streptavidine HRP pour une expérience de photocrosslinking sur cellules entières.....	- 131 -
Figure 59 : Image d'un western blot réalisé avec de la streptavidine HRP pour des milieux de photocrosslinking sur cellules dialysés ou purifiés sur billes . ....	- 132 -
Figure 60 : Spectre MALDI-TOF d'un digest de lysat cellulaire (cellules photoirradiées en présence de photo(R/W) <sub>9</sub> ) après purification sur billes de streptavidine.. ....	- 134 -
Figure 61 : Schéma de la stratégie "adaptée" (N°2) de photocross-linking <i>in cellulo</i> .....	- 136 -
Figure 62 : Gel 1D SDS-PAGE coloré au bleu de Coomassie et Western blot avec de la streptavidine HRP pour l'expérience de photocross-linking sur membranes.....	- 137 -
Figure 63 : Gel d'électrophorèse 1D SDS-PAGE coloré au bleu de Coomassie pour les milieux de pull-down + photocross-linking (1), au pull-down simple (2) et au contrôle négatif (3). ....	- 141 -
Figure 64 : Présentation des séquences en acides aminés des protéines EWS, FLI1 et EWS-FLI1 et de leur degré d'homologie. ....	- 147 -
Figure 65 : Schéma de la perturbation de l'interaction de la protéine EWS-FLI1 avec l'ARN hélicase conduisant à l'inhibition de l'activité de transcription du complexe. ....	- 148 -
Figure 66 : Spectres MALDI-TOF pour la gamme de m/z 1100-2500 montrant le taux d'incorporation des acides aminés marqués <sup>13</sup> C <sub>6</sub> Lys et <sup>15</sup> N <sub>4</sub> Arg seuls et en mélange 1-1 des cellules (L) et (H)...	- 152 -
Figure 67 : Figure illustrant la vérification du taux d'incorporation des acides aminés marqués dans les protéines.....	- 153 -
Figure 68 : Graphique de type volcano plot obtenu pour deux réplicats biologiques de l'expérience contrôle EF (H) / EF (L).....	- 154 -
Figure 69 : Graphique de type volcano plot pour la comparaison du protéome des cellules EF (H) et des cellules 3T3 (L).. ....	- 156 -
Figure 70 : Réseau d'interactions protéine-protéine généré sur le logiciel STRING à partir de la liste de protéines surexprimées dans les cellules EF par rapport aux cellules 3T3.. ....	- 157 -
Figure 71 : Nœud du réseau d'interaction correspondant au processus de glycolyse .....	- 158 -

Figure 72 : Graphique de type volcano plot pour la comparaison du protéome des cellules EF (H) et des cellules EF (L) incubées avec le peptide photo(R/W) <sub>9</sub> (18 h, 5 µM).....	- 163 -
Figure 73 : Modifications du métabolisme caractéristiques de la transformation tumorale .....	- 164 -
Tableau 1 : Etiquettes communément utilisées en purification d'affinité .....	- 47 -
Tableau 2 : Exemple de détergents classiquement utilisés en biochimie et propriétés associées....	- 53 -
Tableau 3 : Trois groupements photoactivables communément utilisés : la benzophénone (Bp), le phényle azide (Pa) et le trifluorométhyl phényl diazirine (TFMFD). .....	- 57 -
Tableau 4 : Avantages et inconvénients des différentes techniques d'étude des interactions protéine-protéine. ....	- 75 -
Tableau 5 : Origines et séquences des CPPs les plus étudiés (adapté de [207]).....	- 80 -
Tableau 6 : Exemples de fluorophores communément utilisés en immunofluorescence.....	- 104 -
Tableau 7 : Attribution des doublets observés dans le spectre MALDI-TOF présenté (Figure 42).....	- 109 -
Tableau 8 : Attribution des fragments du spectre MALDI-TOF/TOF présenté Figure 43.....	- 111 -
Tableau 9 : Tableau récapitulatif des observations et comparaisons réalisées pour les modes de fragmentations CID, haute et basse énergie, et ETD des ions photocross-linkés 2803-2813.....	- 124 -
Tableau 10 : Présentation des différents échantillons de l'expérience de pull-down associée au photocross-linking .....	- 140 -
Tableau 11 : Nombres moyens de peptides et de protéines identifiés pour les différentes pistes pour les deux réplicats biologiques.. .....	- 142 -
Tableau 12 : Extrait du tableau de résultats obtenu avec le logiciel SAINT pour l'expérience de Pull-down simple.. .....	- 146 -
Tableau 13 : Récapitulatif des protéines de liaisons de l'actine figurant dans la liste des protéines significativement sous-exprimées dans les cellules tumorales EF par rapport aux cellules non-tumorales 3T3.....	- 161 -

## **Figures de l'annexe instrumentale : Spectrométrie de masse**

Figure annexe 1: Structure d'un spectromètre de masse. ....	- 207 -
Figure annexe 2 : Schéma de principe de la production d'ions par MALDI .....	- 208 -
Figure annexe 3 : Schéma du processus d'ionisation par électronébulisation. ....	- 209 -
Figure annexe 4 : Principe de la séparation des ions et de la mesure de leur rapport m/z avec un analyseur à temps de vol (TOF). ....	- 210 -
Figure annexe 5 : Schéma de la réflexion électrostatique des ions dans un analyseur TOF en mode réflecteur .....	- 210 -
Figure annexe 6 : A = Schéma du LTQ B = Détection des ions piégés .....	- 211 -
Figure annexe 7 : Vue en coupe de l'analyseur Orbitrap. Les ions sont introduits dans l'analyseur au niveau de la flèche rouge. ....	- 211 -
Figure annexe 8 : Représentation schématique du spectromètre de masse MALDI-TOF/TOF 4700 Proteomics Analyzer.....	- 212 -
Figure annexe 9 : Schéma du spectromètre de masse LTQ-Orbitrap XL (Thermo Scientific) .....	- 213 -

**Listes des annexes diverses (= Hors annexe instrumentale spectrométrie de masse)**

<b>Annexe II 1:</b> Recherches des peptides photocross-linkés (R/W) <sub>9</sub> -BSA avec Mascot sur P.D 1.3 ....	- 215 -
<b>Annexe II 2:</b> Spectres de fragmentation CID (basse et haute énergie) et ETD des ions de (M+H) <sup>+</sup> 2803-2813.....	- 217 -
<b>Annexe II 3 :</b> COMPARAISON A : Pull-down + photocross-linking versus contrôle négatif .....	- 222 -
<b>Annexe II 4 :</b> COMPARAISON B : Pull-down versus contrôle négatif .....	- 226 -
<b>Annexe II 5:</b> Attributions des spectres MS/MS aux peptides [269-304] et [373-386] de la protéine. ....	- 229 -
<b>Annexe II 6:</b> Liste de protéines sur- ou sousexprimées dans les cellules EF (H) par rapport aux cellules 3T3.....	- 230 -
<b>Annexe II 7:</b> Réseaux d'interactions des protéines sous-exprimées.....	- 234 -
<b>Annexe II 8 :</b> Volcano plots obtenus pour les expériences de protéomique différentielle avec le marquage SILAC.....	- 235 -





# Préface

Les peptides vecteurs ou CPPs, courtes séquences peptidiques (< 30 acides aminés), ont la capacité de traverser la membrane plasmique des cellules avec une toxicité très limitée. Depuis leur découverte il y a aujourd'hui bientôt trois décennies, ils ont suscité beaucoup d'intérêt dans la communauté scientifique du fait de leur aptitude à transporter avec eux des molécules, appelées cargos, de tailles parfois très importantes par rapport à celles des CPPs. Alors considérés comme des molécules inertes, les CPPs ouvraient des perspectives très intéressantes en termes de vectorisation. Cependant, depuis quelques années, de plus en plus d'études remettent en cause l'absence d'activité biologique des CPPs et le terme de bioportide a même été créé pour désigner les **CPPs ayant un effet biologique intrinsèque** [1]. Parmi eux, se trouve le peptide (R/W)<sub>9</sub> de séquence RRWWRRWRR, peptide vecteur largement étudié au Laboratoire des BioMolécules (UMR 7203) dans lequel j'ai réalisé mes travaux de thèse. Il a été montré au cours de précédents travaux [2], que ce peptide était capable de modifier le phénotype de cellules tumorales de type EF, modèle cellulaire du sarcome d'Ewing, caractérisé par un cytosquelette d'actine très désorganisé. Au bout de quelques heures d'incubation avec le peptide vecteur (R/W)<sub>9</sub>, un changement de morphologie est observé ainsi qu'un retour des fibres de stress (assemblage de microfilaments d'actine en faisceaux). Ce changement de morphologie, **lié à un remodelage du cytosquelette d'actine**, s'accompagne d'une diminution de la vitesse de migration et d'une diminution de la capacité à croître en indépendance d'ancrage, toutes deux étant caractéristiques des cellules cancéreuses.

De manière générale, il est nécessaire de comprendre comment ces CPPs peuvent affecter la dynamique des cellules de manière à pouvoir un jour les utiliser en clinique sans craindre d'effets secondaires liés à leur activité biologique intrinsèque plutôt qu'à celle de la molécule vectorisée. Dans le cas du peptide (R/W)<sub>9</sub>, cette étude est d'autant plus intéressante du fait de son activité de remodelage du cytosquelette d'actine observé sur les cellules tumorales EF. En effet, l'étude du mode d'action du peptide (R/W)<sub>9</sub> est susceptible d'apporter des éléments qui pourront faire évoluer la compréhension et le traitement du sarcome d'Ewing, cancer très agressif des os et des tissus conjonctifs, touchant majoritairement les adolescents et ayant un très mauvais pronostic en raison de son fort potentiel métastatique.

L'objet de ce travail de thèse a ainsi été d'avancer dans la **compréhension de l'effet moléculaire observé du peptide vecteur (R/W)<sub>9</sub> sur le phénotype de cellules tumorales EF**. Le terme d'« effet moléculaire » fait référence à l'ensemble des interactions protéine-protéine ou peptide-protéine potentiellement impliquées dans le mécanisme d'action du peptide (R/W)<sub>9</sub>.

Le sujet des **interactions protéine-protéine** se trouve par conséquent au cœur de cette étude et j'ai ainsi choisi d'y consacrer une part importante de l'introduction. Dans un premier chapitre, les interactions protéine-protéine, impliquées dans tous les mécanismes de la vie d'une cellule, seront présentées. L'exemple des interactions protéine-protéine dans la régulation de la dynamique de l'actine sera traité en détail de manière à pouvoir interpréter nos résultats et discuter par la suite l'effet du CPP (R/W)<sub>9</sub> sur le remodelage du cytosquelette. Dans le deuxième chapitre, différentes méthodes pour l'étude des interactions protéine-protéine seront abordées, dont un certain nombre

qui ont été utilisées au cours de ma thèse. Puis, le troisième et dernier chapitre de l'introduction sera consacré aux CPPs, en particulier à leur(s) mode(s) d'internalisation et à leur devenir intracellulaire.

La partie résultats sera, elle, structurée en quatre chapitres. Tout d'abord, un chapitre concernant le travail réalisé *in vitro* pour caractériser l'interaction du CPP (R/W)<sub>9</sub> avec l'actine par cross-linking chimique et spectrométrie de masse. Puis, les trois chapitres suivants correspondent aux différentes stratégies développées pour identifier des interactions peptide-protéine ou protéine-protéine *in cellulo* ou dans des lysats cellulaires susceptibles d'expliquer l'effet du CPP (R/W)<sub>9</sub>. Il s'agit de **stratégies ciblées** développées pour la **capture des partenaires d'interaction du peptide**, (photocross-linking *in cellulo* (chapitre 5) et purification d'affinité couplée à la spectrométrie de masse (chapitre 6)) et d'une **stratégie indirecte** correspondant à une approche de protéomique différentielle ayant pour objectif d'évaluer **l'impact du peptide (R/W)<sub>9</sub> sur l'expression des protéines** des cellules EF pour en déduire son mode d'action (Chapitre 7).

# Chapitre 1 : Des interactions protéine-protéine au phénotype cellulaire, l'exemple du cytosquelette d'actine

La protéomique est la discipline qui étudie l'ensemble des protéines et leurs rôles dans une cellule, un organe, un tissu, un organe ou un organisme à un moment donné et sous des conditions données. Dans un premier temps, la protéomique s'est attachée à l'identification globale des protéines extraites d'une culture cellulaire, d'un tissu ou d'un fluide biologique et à la détermination de leur localisation dans les compartiments cellulaires, de leurs éventuelles modifications post-traductionnelles ainsi que de leur quantité. Aujourd'hui, elle s'oriente de plus en plus vers la compréhension du rôle de ces protéines de façon intégrée dans la cellule et vers l'évaluation de l'incidence des modifications de leur expression au niveau d'un organisme.

Jusqu'au début des années 1990, notamment en raison du manque d'outils appropriés, les fonctions des protéines ont été étudiées de manière indépendante. Cependant, les fonctions cellulaires étant rarement attribuées à une seule molécule, il est apparu nécessaire de prendre en compte les interactions existant entre différentes protéines organisées en modules fonctionnels [3] pour comprendre le rôle d'une protéine en particulier. De plus, la plupart des protéines sont impliquées dans plusieurs processus cellulaires distincts et en association avec différentes protéines. La complexité des interactions est également augmentée en raison de la diversité spatiale et temporelle des protéines. En effet, une même protéine peut faire partie de complexes protéiques différents en fonction de sa localisation subcellulaire ou du stade du cycle cellulaire. Les modifications post-traductionnelles peuvent également influencer sur la capacité d'une protéine à établir une interaction. Finalement les interactions protéine-protéine varient énormément en termes de nature, de force et de durée.

Ainsi la compréhension de la fonction cellulaire d'une protéine peut être une tâche extrêmement lourde en raison de l'extraordinaire complexité des interactions protéine-molécule de façon générale, et plus particulièrement protéine-protéine dans lesquelles elle est susceptible d'être impliquée.

Dans ce premier chapitre, nous allons commencer par poser quelques bases pour la compréhension de la « sociologie moléculaire de la cellule » [4]. Ensuite, et compte tenu du sujet développé, nous nous intéresserons à un des réseaux d'interactions protéine-protéine les plus étudiés, à savoir le cytosquelette d'actine et ses protéines de liaison (ABPs pour *actin binding proteins*). Finalement nous introduirons le contexte biologique de l'étude, le sarcome d'Ewing, essentiellement caractérisé par l'expression d'une protéine de fusion EWS-FLI1 induisant d'importantes modifications du phénotype des cellules en particulier au niveau de leur cytosquelette.

## **I. Au sujet des interactions protéine-protéine**

### **1. Les différents types d'interactions protéine-protéine**

Hormis quelques rares exemples tel que l'attachement covalent de l'ubiquitine ou des protéines SUMO aux protéines de signalisation [5], les protéines se lient les unes aux autres très majoritairement de manière non-covalente. Ces liaisons non-covalentes, dont l'énergie est comprise entre 0 et 40 kJ.mol<sup>-1</sup>, résultent de la combinaison de liaisons hydrophobes, de liaisons hydrogènes, d'interactions de Van der Waals et/ou de liaisons ioniques, plus fortes, au niveau de domaines spécifiques de chacune des deux protéines. Ces domaines d'interactions peuvent être limités à de petites poches ou s'étendre sur de plus grandes surfaces et peuvent ainsi impliquer de quelques résidus à plusieurs centaines d'acides aminés. Il est important de comprendre la nature des interactions protéine-protéine car les différents types de liaison vont être plus ou moins sensibles aux changements de l'environnement cellulaire, que ce soit la composition ionique du milieu, le pH ou d'autres paramètres chimiques.

On distingue généralement les **interactions protéiques stables** des **interactions transitoires**.

Les **interactions stables** sont celles impliquées dans l'association de protéines qui peuvent être purifiées sous forme de complexes constitués de plusieurs sous-unités de même nature ou de natures différentes. C'est le cas par exemple des protéines du ribosome qui n'existent dans les cellules que sous forme associées.

Ce sont les **interactions transitoires** qui sont vraisemblablement à l'origine de la majorité des fonctions cellulaires. Comme leur nom l'indique, les interactions transitoires sont limitées dans le temps et conditionnées par différents éléments tels qu'une modification post-traductionnelle, un changement conformationnel ou l'arrivée dans un compartiment cellulaire particulier. Par exemple, ce sont des interactions transitoires qui sont à l'origine de la transmission de signaux intracellulaires telle que la liaison d'un ligand à un récepteur couplé à la protéine G. Les complexes résultant d'interactions transitoires sont beaucoup plus compliqués à étudier car l'identification des protéines et/ou des conditions nécessaires à l'observation de ces complexes sont des prérequis à leur étude. Il faudrait pour cela aller vers l'exhaustivité dans l'identification des éléments d'un système, ce qui constitue un des défis parmi les plus importants.

### **2. Les domaines d'interactions des protéines**

Les interactions protéine-protéine reposent sur l'existence de **domaines d'interaction** définis au sein des protéines. Ces domaines d'interaction peuvent permettre de définir la localisation de la protéine dans la cellule, permettre de reconnaître les modifications post-traductionnelles des protéines partenaires, être à l'origine de la formation d'un complexe multiprotéique ou encore contrôler la conformation, l'activité ou la spécificité pour un substrat d'un enzyme.

Typiquement, un domaine d'interaction d'une protéine reconnaît un **élément central déterminant ainsi que des résidus adjacents contigus ou non qui permettent des contacts supplémentaires** et améliorent la spécificité de l'interaction. Par exemple, le domaine SH3 de la tyrosine kinase cytoplasmique (Csk) reconnaît un motif central PXXP (P : proline, X : n'importe quel acide aminé) de l'extrémité de la PEP tyrosine phosphatase qui adopte alors une hélice polyproline de type II typique des sites de liaison aux domaines SH3. Ensuite, le domaine SH3 Csk est également en interaction avec

deux résidus hydrophobes de la protéine PEP grâce à une poche de liaison distincte qui assure la spécificité de l'association *in vivo* [6].

Cependant, dans certains cas, **l'affinité d'un seul domaine d'une protéine pour un motif peptidique peut être suffisante** pour obtenir une interaction spécifique.

Les domaines d'interaction sont caractérisés par une séquence consensus d'acides aminés. De nombreux domaines sont présents en plusieurs centaines de copies dans le protéome et sont utilisés pour réguler différentes fonctions cellulaires. Certains domaines ont des fonctions bien spécifiques comme le domaine SH2 qui est dédié à la signalisation des tyrosines kinases [7]. Au contraire, d'autres domaines sont communs à un grand nombre de protéines et impliqués dans des fonctions cellulaires variées. C'est le cas du domaine SH3 qui régule à la fois la transduction de signaux, le trafic de protéines et de vésicules, l'organisation du cytosquelette ou encore la biogénèse d'organelles. La cellule fait ainsi appel à un ensemble de domaines d'interaction qui peuvent être combinés de différentes manières pour permettre l'action de ses systèmes de régulation [8].

C'est sur la base des connaissances acquises au fil des années concernant ces domaines d'interaction qu'a été développée la base de données Pfam (*Protein families database*) qui regroupe une large collection d'alignement de séquences ainsi que des annotations fonctionnelles et des références bibliographiques de chaque famille de protéines. Elle a pour but de permettre de déterminer la fonction de nouvelles protéines en se basant sur l'existence dans leur séquence de domaines caractéristiques d'une famille particulière de protéines [9]. La fonction prévue par cette approche bioinformatique doit ensuite être confirmée par des expériences *in vitro* mais constitue un point de départ intéressant.

### 3. Régulation et effets biologiques des interactions protéine-protéine

Beaucoup des protéines impliquées dans les voies de signalisation ont une activité enzymatique et ont la capacité de **modifier leur substrat de manière post-traductionnelle** [10]. La phosphorylation d'une protéine est la modification post-traductionnelle la plus courante et la mieux comprise pour la régulation d'interactions protéine-protéine mais d'autres modifications comme l'acétylation et l'hydroxylation sont vraisemblablement également impliquées dans la modulation de ces interactions. Les modifications post-traductionnelles peuvent affecter les interactions protéine-protéine selon deux mécanismes. Premièrement, les modifications post-traductionnelles **changent les propriétés chimiques de l'acide aminé modifié** entraînant la création ou la disparition d'une liaison chimique avec la protéine partenaire. Deuxièmement, les modifications post-traductionnelles peuvent **affecter le repliement de la protéine** et dans le même temps la structure secondaire des domaines d'interaction. Ainsi les modifications post-traductionnelles peuvent soit permettre d'exposer un nouveau domaine à la surface d'une protéine ou, au contraire, rendre un domaine d'interaction inaccessible pour la liaison avec une protéine partenaire [11]. Les protéines 14-3-3 qui sont capables de reconnaître les acides aminés phosphorylés de séquences particulières sont un bon exemple de la régulation d'interactions protéine-protéine par les modifications post-traductionnelles [12].

Phizicky E. M. et Fields S. ont définis cinq catégories d'effets biologiques liés aux interactions protéine-protéine [13]. (1) Dans le cas de protéines allostériques, ces interactions peuvent modifier les propriétés cinétiques de la protéine ainsi que leur l'affinité pour un substrat [14]. (2) Un certain

nombre d'interactions protéine-protéine sont connues pour permettre le passage d'un métabolite intermédiaire directement d'une enzyme à une autre sans libération en solution (*substrate channeling*). (3) Certaines interactions protéine-protéine peuvent donner naissance à de nouveaux sites de fixation. (4) Des interactions protéine-protéine conduisent à l'inactivation d'une des deux protéines, la protéine partenaire est alors qualifiée d'inhibiteur. (5) Des interactions protéine-protéine peuvent modifier la spécificité d'une protéine pour son substrat.

### 4. Les réseaux d'interactions des protéines

Les interactions protéine-protéine ne doivent pas être envisagées comme des événements isolés. Ces interactions doivent au contraire être replacées au sein de réseaux denses et interconnectés. En effet, l'activation d'une cascade de signalisation reposant sur un complexe protéique implique également bien souvent des effecteurs placés à la fois en amont et en aval de cette voie de signalisation. Il est important également de garder à l'esprit qu'une même protéine peut être impliquée dans différentes interactions protéine-protéine.

**La cartographie systématique des interactions protéine-protéine ou cartographie de l'interactome** a été initiée pour des organismes modèles avec des processus biologiques définis [15], [16] avant d'être étendue à l'échelle du protéome. Si l'étude des interactions protéine-protéine peut permettre de comprendre le rôle d'une protéine ; la compréhension de l'organisation globale d'un réseau d'interaction peut, elle, permettre d'améliorer la connaissance de mécanismes complexes en biologie fondamentale et d'expliquer le développement de différentes pathologies. Ainsi, par exemples, un réseau à 186 interactions impliquées dans la maladie d'Huntington a été établi [17] tandis que 755 interactions ont été rassemblées dans la voie de signalisation de TGF- $\beta$  (*transforming growth factor- $\beta$* ) [18].

De nombreuses approches ayant pour but d'étudier les fonctions d'une protéine reposent par conséquent sur l'identification de partenaires d'interaction de manière à pouvoir replacer la protéine étudiée dans un réseau d'interactions impliqué dans la régulation d'un processus connu. Nous aborderons ces différentes approches dans le chapitre 2.

## **II. Le cytosquelette d'actine : illustration de la complexité des réseaux d'interactions**

Le cytosquelette interne des cellules eucaryotes est composé de microfilaments d'actine, de microtubules et de filaments intermédiaires. Le cytosquelette est un ensemble dynamique qui joue un rôle crucial dans l'activité cellulaire en contrôlant notamment la morphologie de la cellule, sa motilité et sa capacité à se diviser. Il est également fortement impliqué dans le transport intracellulaire. Nous allons nous intéresser tout particulièrement à la régulation de la dynamique des microfilaments d'actine.

### **1. L'actine : fiche d'identité [19]**

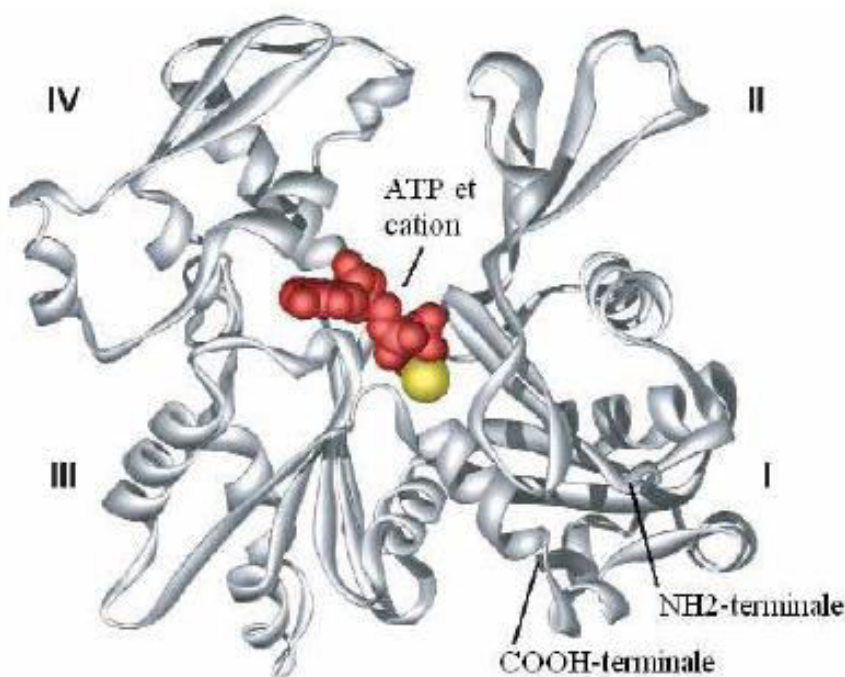
L'actine est une protéine ubiquitaire dans les cellules eucaryotes, très conservée d'un organisme à un autre, et qui présente trois isoformes principales  $\alpha$ -,  $\beta$ -,  $\gamma$ -. C'est dans les cellules musculaires que la concentration totale d'actine est la plus forte ( $\approx 20$  % de la quantité totale de protéines), tandis qu'elle représente en moyenne environ 5 % des protéines dans les autres types de cellules où elle est impliquée dans la morphologie cellulaire, le transport d'organelles, la régulation du transport des ions ou la réponse à des signaux extracellulaires transmis par les récepteurs.

L'actine se trouve majoritairement dans le cytoplasme des cellules. L'actine monomérique (Actine G pour actine globulaire) s'assemble dans les cellules pour former des filaments d'actine (Actine F) de structure légèrement hélicoïdale de 5 à 9 nm de diamètre (Figure 2).

La détermination de la structure atomique de l'actine G (complexée avec la ADNase I [20]) est à la base de la compréhension actuelle de l'actine. L'actine G est constituée de 375 acides aminés et a une masse moléculaire de 43 kDa. Elle est repliée de manière à former deux grands domaines qui eux-mêmes comprennent deux sous-domaines (les sous-domaines sont numérotés de I à IV selon la nomenclature de Kabsch[20]).

Dans cette structure de l'actine représentée figure 1, on note la présence de deux lobes constitués chacun de deux sous-domaines. Le premier sous-domaine (sous-domaine I) est constitué par les segments non-consécutifs 1-32, 70-144 et 338-372 et contient donc les deux extrémités de la protéine ( $\text{NH}_2$ -terminale et  $\text{COOH}$ -terminale). Le sous-domaine II qui complète le premier lobe est formé par le segment 33-69. L'autre lobe comporte les sous-domaines III, constitué des segments 145-180 et 230-337, et IV qui englobe la partie centrale de la protéine entre les résidus 181 et 269. Au coeur de ce globule, dans le sillon formé entre les deux domaines, se trouve un site de complexation d'un nucléotide (ATP ou ADP) associé à un cation divalent ( $\text{Mg}^{2+}$  ou  $\text{Ca}^{2+}$ ) [21]–[23]. Cette association de l'actine avec ces cofacteurs essentiels permet la stabilisation de la protéine et limite donc sa dénaturation.

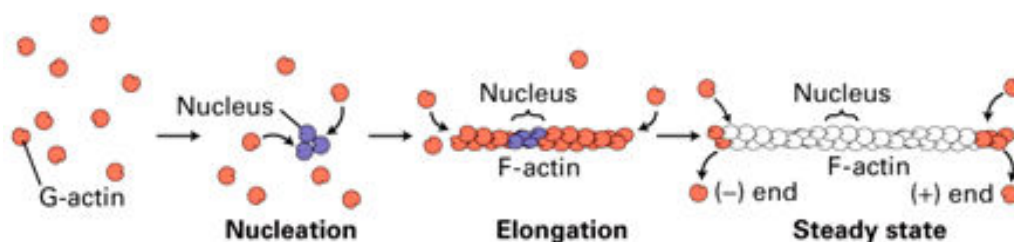




**Figure 1 : Représentation schématique de la structure cristallographique d'un monomère d'actine.** Les 4 sous-domaines (I à IV) sont représentés en gris, l'ATP en rouge et le cation divalent en jaune. (Figure d'après [24])

## 2. Polymérisation de l'actine

*In vitro*, en présence de 100 mM de KCl, d'ATP et de  $MgCl_2$ , les monomères d'actine G purifiés et associés à une molécule d'ATP polymérisent pour former des filaments d'actine F. Ce processus se déroule en trois étapes théoriques schématisées dans la figure 2 : la nucléation, l'élongation et l'état stationnaire [25].



**Figure 2 : Schéma de la polymérisation de l'actine** (Figure d'après [24]).

Au cours de la nucléation, les monomères d'actine s'agregent en dimères puis en trimères instables jusqu'à former des noyaux de polymérisation, les *nuclei* qui sont les premiers oligomères et dont la dissociation est moins favorable que l'association avec d'autres monomères. Après la formation de ces *nuclei* intervient l'étape d'élongation au cours de laquelle des monomères d'actine sont additionnés au niveau des deux extrémités du filament. Lors de l'incorporation d'un monomère d'actine-ATP dans le polymère, le nucléotide est hydrolysé pour donner de l'ADP avec libération d'un phosphate inorganique ( $P_i$ ) [26]. L'hydrolyse du nucléotide s'accompagne de réarrangements conformationnels qui permettent la différenciation entre les actines présentes aux deux extrémités



du filament (polarisation du filament). Les cinétiques de polymérisation aux deux extrémités ne sont pas les mêmes et la vitesse de polymérisation de l'extrémité (+) dite barbée est 5 à 10 fois plus rapide que celle à l'extrémité (-) dite pointue [27].

De façon générale, cette étape voit la part de l'actine F augmenter alors que la concentration en monomère diminue [28]. Lorsque celle-ci atteint la concentration critique de polymérisation de l'actine ( $C_c$ ), le régime devient alors équilibré. On observe simultanément une polymérisation d'actine à l'extrémité barbée et une dépolymérisation au niveau de l'extrémité pointue. L'état d'équilibre se caractérise donc par une égalisation des vitesses d'assemblage à l'extrémité (+) et de désassemblage à l'extrémité (-). Pour boucler ce phénomène d'échange de monomères ou *treadmiling*, le nucléotide ADP est remplacé par un ATP afin de restaurer le monomère polymérisable actine-ATP.

### *In vivo*

*In vivo*, l'étude de la dynamique du cytosquelette d'actine devient beaucoup plus complexe. La polymérisation des monomères d'actine G ne repose plus seulement sur la concentration en cofacteurs et sur le ratio actine G/actine F mais également, comme nous allons le voir dans la suite, sur un nombre très important d'interactions protéine-protéine qui sont régulées en fonction des contraintes auxquelles les cellules sont soumises.

## 3. Les protéines de liaison de l'actine et leurs rôles

La localisation, la longueur, la polarisation des filaments d'actine ainsi que l'aspect temporel de l'association des monomères d'actine pour former ces filaments sont régulés par un grand nombre de protéines de liaison de l'actine (*actin binding proteins* ou ABPs) qui sont, elles, régulées par les cascades de signalisation initiées à la surface de la cellule. Les filaments sont à leurs tours organisés de manière tridimensionnelle grâce à un autre groupe de protéines de liaison qui peuvent s'associer à ces différents filaments, les séquestrer, donner naissance à des faisceaux ou encore les ancrer à la membrane.

L'actine est une des protéines, si ce n'est la protéine, impliquée dans le plus grand nombre d'interactions protéine-protéine dans la cellule. Ainsi, il a été montré jusqu'à présent que plus de 150 protéines contiennent un domaine de liaison à l'actine et sont susceptibles d'influencer son état de polymérisation [29]. Il est ainsi hors de question de dresser la liste de tous ses partenaires et de leur action. Nous allons simplement essayer de définir les rôles principaux et de donner des exemples significatifs de protéines impliquées dans ces fonctions.

### a. La régulation de la polymérisation des microfilaments

#### **Régulation de la concentration en actine G libre**

Dans les cellules, une grande quantité d'actine est maintenue sous forme monomérique (actine G) par des protéines dites de **séquestration** [30]. La séquestration de l'actine G par ces protéines permet de conserver un stock important d'actine G dans la cellule tout en maintenant la concentration d'actine G libre basse dans la cellule lorsque la polymérisation des microfilaments a atteint l'état stationnaire. Cependant, suite à une stimulation appropriée, la polymérisation de

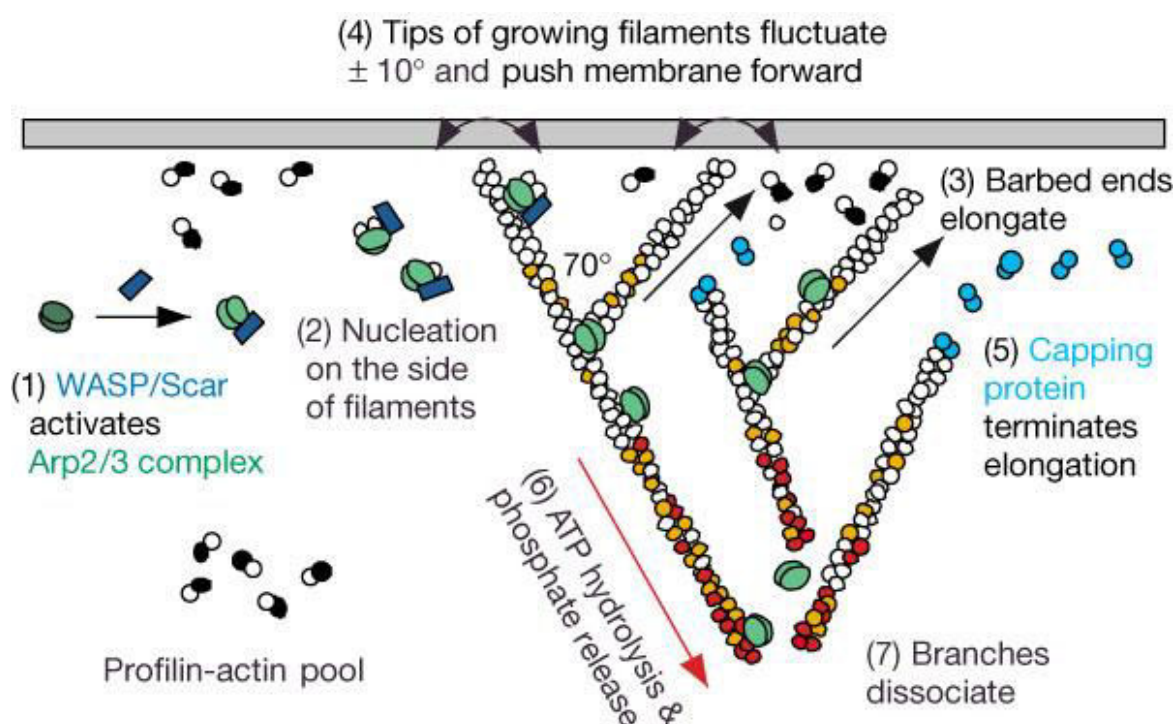
l'actine de manière rapide et locale est rendue possible par la libération quasi-instantanée des monomères séquestrés.

Il apparaît que deux protéines de liaison de l'actine G sont impliquées dans la séquestration de l'actine : La thymosine  $\beta 4$ , une petite protéine de 5 kDa, qui a été identifiée comme l'agent de séquestration principal de l'actine [31] et la profiline, une protéine de masse moléculaire comprise entre 12 et 15 kDa découverte en 1977 [32] et connue pour former un complexe 1:1 avec l'actine G.

### Nucléation et régulation de la nucléation

Le complexe Arp2/3 est un complexe de sept sous unités dont les protéines Arp2 et 3 (*Actin related protein*) et cinq autres protéines nommées p41, p34, p2, p20 et p16 [33]. Ce complexe induit la **nucléation** de nouveaux filaments après stimulation par les protéines de la famille WASP (*Wiskott-Aldrich Syndrom Protein*)/Scar [34] qui possèdent un domaine WA dans leur partie C-terminale. La région "verproline homology", dite W, recrute deux monomères d'actine et la séquence acide, dite A, active le complexe Arp2/3.

Le complexe Arp 2/3 est le point de départ du pontage de filaments nouvellement formés et organise les filaments d'actine en un réseau ayant une structure en Y [35] comme schématisé dans la figure 3.



**Figure 3 : Représentation schématique de la dynamique de polymérisation de l'actine.** L'actine-ATP est représentée en blanc, l'actine ADP-P en orange, l'actine-ADP en rouge et la profiline en noir (Figure d'après [36]).

### Régulation des échanges de monomères

Les **protéines de coiffe** telles que CapZ ( $=\beta$ -actinine, 60 kDa) contrôlent la longueur et la vitesse de croissance des filaments en se liant avec une très haute affinité à leur extrémité barbée (+) [37] et en empêchant ainsi tout échange de monomères.

La gelsoline (87 kDa) dont l'activité est dépendante du pH, de la présence de  $\text{Ca}^{2+}$  et de la liaison aux phosphoinositides peut lier l'actine F et induire la coupure des filaments en deux en affaiblissant les liaisons non covalentes entre les monomères d'actine [38]. Ensuite la gelsoline reste attachée à l'extrémité barbée du filament empêchant son élongation [39] ou son attachement à un autre filament. La gelsoline joue ainsi à la fois un rôle de **protéine de ségrégation et de coiffe**.

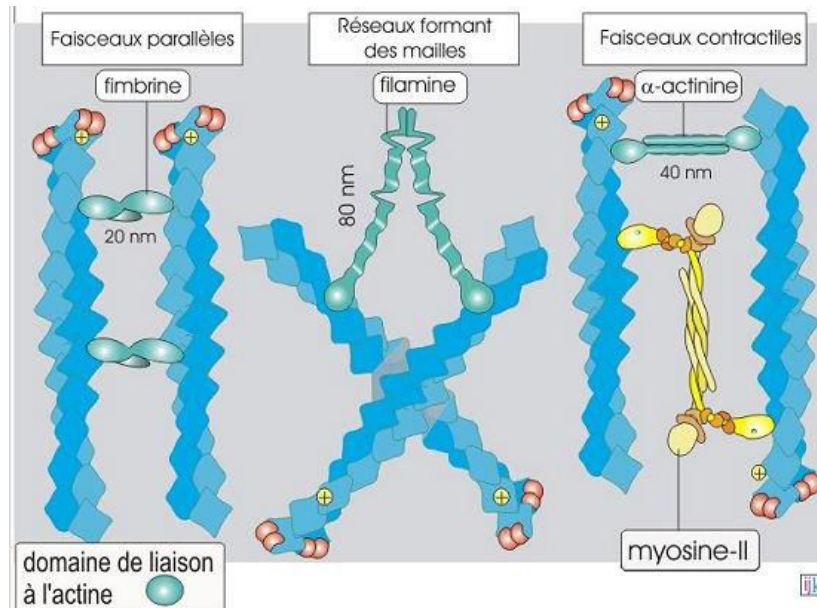
Les protéines de la famille ADF (*actin depolymerizing factor*)/Cofiline sont des protéines de 15 à 18 kDa qui freinent la croissance de l'extrémité barbée (+) des filaments d'actine par une activité de **dépolymérisation** [40]. Ces protéines sont activées par une déphosphorylation d'une sérine de leur partie N-terminale.

### Stabilisation latérale des filaments

La tropomyosine, la nébuline et la caldesmone sont des protéines allongées qui se lient sur les côtés des filaments d'actine et qui contribuent à la **stabilisation** de ces filaments. La tropomyosine augmente la force de tension des filaments d'actine et les protège de l'action de dépolymérisation de l'ADF [41]. La nébuline est très grande protéine (600-900 kDa) qui peut lier jusqu'à 200 monomères d'actine et qui semble impliquée dans la régulation de la longueur des filaments d'actine dans les muscles striés [42]. La fonction de la caldesmone est moins connue, mais un ensemble d'éléments semble indiquer qu'avec la tropomyosine elle forme un complexe sensible au  $\text{Ca}^{2+}$  et à la calmoduline, régulateur des interactions actine-myosine dans les muscles lisses ainsi que dans les cellules non musculaires [43]. La caldesmone est phosphorylée par des tyrosines kinases intervenant dans le cycle cellulaire et pourrait jouer un rôle dans la réorganisation des filaments d'actine au cours de la mitose [44].

### b. Assemblage des microfilaments

Les microfilaments d'actine s'assemblent pour donner deux types de structures : des réseaux et des faisceaux. On distingue ensuite les faisceaux parallèles des faisceaux anti-parallèles ou contractiles soit trois arrangements au total, schématisés Figure 4.



**Figure 4 : Les trois arrangements de l'actine** (figure tirée du site internet [www.ulysses.u-bordeaux.fr](http://www.ulysses.u-bordeaux.fr))

Bien que les mécanismes d'assemblage des microfilaments n'aient pas encore été complètement élucidés, il semblerait qu'ils dépendent de l'expression d'un mélange approprié de protéines de liaison à l'actine, qui sont un prérequis à l'auto-assemblage de structures particulières, et de signaux médiés par des petites GTPases de la famille Rho (Rho, Rac, Cdc42) [45].

### Réseaux d'actine

Les réseaux sont formés grâce à des protéines de **réticulation**, qui associent les microfilaments de manière orthogonale. Un bon exemple de ce type de protéines est la **filamine**, protéine homodimérique avec deux grandes sous-unités de 280 kDa qui présente une forme de V avec chacune des branches qui peut s'associer à un microfilament [46]. Le réseau tridimensionnel de filaments d'actine joue un rôle fondamental dans le maintien de la forme cellulaire. Chez toutes les cellules, un réseau d'actine particulièrement important, l'actine corticale juxta-membranaire se trouve juste sous la membrane plasmique. En plus de son **rôle dans la morphologie cellulaire**, ce réseau permet le **bon déroulement des mécanismes d'endocytose et d'exocytose**. Ce réseau d'actine corticale juxta-membranaire est en général structuré par des molécules de la famille de la spectrine. La spectrine est un tétramère (1050 kDa) qui s'associe d'une part par ses extrémités avec l'actine et qui interagit d'autre part avec différentes protéines de la membrane. La spectrine permet donc l'ancrage d'un réseau d'actine à la membrane. Ce réseau améliore la résistance mécanique de la membrane et permet également de limiter le déplacement des protéines dans le plan de la membrane [47].

### Faisceaux d'actine

Les faisceaux d'actine sont liés à la présence de protéines de **pontage**. Ces protéines, en général sous forme de dimère ou de tétramère, possèdent à chacune de leurs extrémités, un domaine de fixation à l'actine. Les filaments d'actine peuvent être assemblés en liasses de façon à ce que tous soient orientés de la même façon, il s'agit alors de **faisceaux parallèles**. Dans le cas contraire d'une association de filaments de polarités opposées on parle de **faisceaux contractiles ou anti-parallèles**.

Les faisceaux parallèles sont présents dans les filopodes et les microvillosités. Les microvillosités sont des replis dans la membrane plasmique de 1 à 2  $\mu\text{m}$  composés de 20 à 30 microfilaments parallèles attachés les uns aux autres essentiellement par les protéines villine [48] et fimbrine [49].

La surexpression de la villine dans les cellules induit l'extension de filopodes existants qui sont des projections de la membrane plasmique formées par des faisceaux internes de filaments d'actine, ainsi que la formation de nouveaux filopodes. Par conséquent le pool de villine et de fimbrine restreint le nombre de faisceaux parallèles, de filopodes et de microvillosités. Dans les cellules contenant ces unités de constructions, l'activation de la GTPase Cdc42 stimule la formation des faisceaux par un mécanisme dépendant de WASP [50], [51]. La relation entre l'apparition d'embranchements par nucléation et l'assemblage de faisceaux de filaments parallèles dans les filopodes est encore mal comprise.

Les faisceaux contractiles tels que les fibres de stress sont structurés par la protéine  $\alpha$ -actinine (200 kDa). Ces faisceaux sont assez lâches ce qui permet à des moteurs moléculaires tels que la myosine II de se glisser entre les filaments pour permettre des contractions [52]. L' $\alpha$ -actinine semble être également impliquée dans la liaison de protéines aux fibres de stress telles que les protéines PDZ-LIM et des protéines kinases. La palladine est une autre protéine de pontage de l'actine qui interagit avec l' $\alpha$ -actinine et est localisée avec la même périodicité que cette dernière le long des fibres de stress [53]. Différentes isoformes de la tropomyosine sont localisées et semblent être importantes pour la maintenance des fibres de stress [54]. L'instabilité des fibres de stress en l'absence de tropomyosine est liée à deux phénomènes : 1) les fibres de stress deviennent plus sensibles à la dépolymérisation par l'ADF/cofiline, 2) l'absence de tropomyosine diminue les contractions liées à la myosine conduisant à la perte d'organisation des fibres de stress [55].

Nous retrouverons un certain nombre de ces protéines jouant un rôle clé dans la dynamique de l'actine dans la partie résultats de ce manuscrit.

#### 4. Le rôle du cytosquelette dans la réponse de la cellule aux contraintes mécaniques externes

Toutes les cellules dépendent de stimulations mécaniques pour leur homéostasie. Les cellules ont la capacité de sentir l'intensité et la direction de forces mécaniques, d'intégrer le signal dans l'espace et le temps et d'en extraire l'information pertinente [56], [57]. Les cellules répondent aux changements mécaniques majoritairement en remodelant leur cytosquelette et/ou en adaptant la synthèse de composés de la matrice extracellulaire [58], [59].

Les cellules d'origines mésenchymateuses comme les fibroblastes, les cellules endothéliales et les cellules musculaires adhèrent à la matrice extracellulaire grâce à des protéines spécialisées dans l'adhésion cellules-matrice [60].

Dans tous les types d'adhésions cellule-matrice, les protéines majoritaires sont les **intégrines**. Ces glycoprotéines transmembranaires hétérodimériques, qui interagissent avec la matrice extracellulaire d'une part et qui sont liées aux filaments d'actine sur la face interne de la membrane d'autre part, sont un élément clé de la **mécanotransduction** [61].

La cellule s'adapte mécaniquement en augmentant sa rigidité suite à l'application de forces **au niveau des intégrines** ou lors d'une rigidification de la matrice extracellulaire [62]. Le stress mécanique peut être induit par la liaison d'un ligand des intégrines tel que la **fibronectine**, un des principaux constituant de la matrice extracellulaire qui se lie au récepteur intégrine  $\alpha 5 \beta 1$  [63] *via* un site consensus comprenant la séquence Arg-Gly-Asp (RGD) [64]. Dans ce cas, les intégrines présentes de manière diffuse dans la membrane vont se rassembler pour former des **points d'adhésion focaux**. Ces changements de la taille des points d'adhésion, de leur composition et de leur position interviennent quelques secondes à quelques minutes après l'application du stress mécanique par le ligand [65]. Cependant, lorsque des forces mécaniques d'intensité comparable sont appliquées aux intégrines au moyen de molécules qui ne lient pas le site de liaison RGD du récepteur, la formation de points d'adhésion focaux n'a pas lieu [66].

La formation des points d'adhésion focaux passe par l'activation des petites Rho GTPase et de leurs effecteurs, la Rho-associated kinase (ROCK) et mDia [67]. ROCK entraîne la formation des points d'adhésion focaux en favorisant la phosphorylation de la chaîne légère de la myosine (*Myosin light chain*, MLC) [68]. Il a également été montré que le remodelage de l'actine suite à l'application d'une contrainte mécanique passe par l'activation de la tyrosine phosphatase SHP2 qui à son tour active la tyrosine kinase Src [69].

De plus, des canaux ioniques sont également activés lors de l'application d'une contrainte mécanique au niveau des intégrines [70]. Leur activation conduit à une augmentation de la quantité de  $\text{Ca}^{2+}$  cytoplasmique. Cet influx de calcium dans le cytoplasme peut moduler les interactions calmoduline/caldesmone et ainsi influencer l'organisation du cytosquelette [71].

### III. Cytosquelette et oncogénèse

#### 1. Généralités [72]

Les cancers (ou tumeurs malignes) sont généralement décrits comme des pathologies résultant de la transformation du matériel génétique de cellules normales suite à des mutations successives et conduisant, au final, à une prolifération incontrôlée des cellules cancéreuses obtenues. Il semblerait cependant que le nombre de mutations nécessaires au processus de tumorigénèse varie d'un cancer à un autre [73].

Hanahan D. et Weinberg R. ont publié en 2000 puis en 2011 des revues visant à rassembler les résultats de la recherche sur le cancer et définissant les caractéristiques fondamentales (hallmarks) communes des cellules cancéreuses [74]. Au départ, six caractéristiques principales ont été décrites :

- 1) Autosuffisance en facteurs de croissance ;
- 2) Insensibilité aux facteurs anti-prolifératifs ;
- 3) Capacité d'échapper à l'apoptose ;
- 4) Capacité de division illimitée ;
- 5) Angiogénèse ;
- 6) Capacité à envahir les tissus et à former des métastases.



Dans l'article de 2011, deux nouvelles caractéristiques qualifiées d'émergentes ont été ajoutées aux précédentes [75] :

- 7) Capacité à détourner le métabolisme énergétique ;
- 8) Capacité à échapper à la surveillance du système immunitaire.

La transformation tumorale qui conduit à l'acquisition de ces capacités s'accompagne de changements importants dans la morphologie des cellules transformées et d'une altération des contacts cellule-cellule et cellule-matrice qui contribuent à augmenter sa motilité.

En effet, **de manière à pouvoir former des métastases** à distance du site de la tumeur primaire et pour migrer dans les tissus adjacents, les cellules cancéreuses acquièrent une motilité plus importante **en modifiant la structure de leur cytosquelette**. Cela passe par la **polymérisation de filaments d'actine** (actine F) en différentes protrusions nommées lamella, lamellipodia, filopodia [76] ainsi que invadopodia [77]. Dans la mesure où les protéines de liaison de l'actine jouent un rôle fondamental dans la régulation de la polymérisation de l'actine, elles ont toujours suscité un intérêt important dans le domaine de la recherche sur le cancer.

Au cours des années, les efforts se sont concentrés sur l'identification de nouveaux biomarqueurs du cytosquelette présentant une bonne corrélation entre leur niveau d'expression et le degré de malignité d'une tumeur. Dans une revue publiée en 2013, Gross S.R. fait le bilan des connaissances sur les mécanismes de régulation de différentes protéines de liaisons de l'actine telles que les protéines Arp2/3, les protéines WASP/WAVE, la fascine et les tropomyosines, et récapitule les corrélations établies entre leurs niveaux d'expression et l'origine de la tumeur [78]. Il met ainsi en évidence l'utilité que peuvent présenter ces marqueurs tout en indiquant les difficultés qu'il y a parfois à trancher sur l'action d'une protéine, par exemple N-wasp, qui est sous-réglée dans le cancer du sein mais qui serait surexprimée dans le cancer de l'œsophage.

Les cellules cancéreuses bien que partageant toutes un certain nombre de caractéristiques de base, présentent également des différences telles que la vitesse de croissance et la capacité de se métastaser en fonction du tissu, du type cellulaire dont elles sont issues.

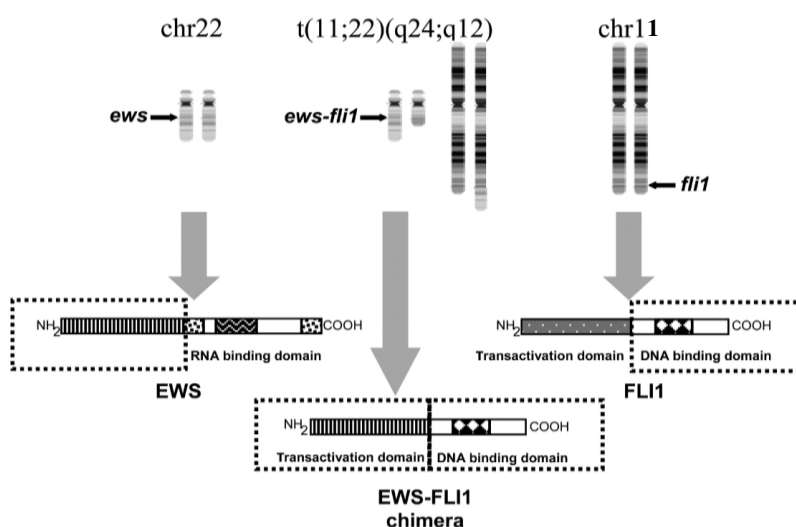
Par la suite nous allons nous focaliser sur le sarcome d'Ewing qui est au cœur de ce projet de thèse. Nous avons travaillé avec les cellules EF, une des lignées cellulaires les plus utilisées pour étudier ce cancer.

## 2. Sarcome d'Ewing : Comment la protéine EWS-FLI1 agit-elle sur le cytosquelette ?

Le sarcome d'Ewing, forme de cancer très agressive avec un potentiel métastatique élevé, a été décrit pour la première fois par James Ewing en 1921 [79]. Un sarcome est un cancer proliférant dans les tissus conjonctifs. Les tumeurs d'Ewing se développent classiquement autour ou dans les os mais peuvent également toucher des tissus mous (sarcome d'Ewing extraosseux) [80]. Les tumeurs d'Ewing, deuxième type de tumeurs des os en terme de fréquence, touchent très majoritairement les adolescents et les jeunes adultes avec la moitié des diagnostics réalisés au cours de la deuxième décennie de la vie [81].

Histologiquement, les sarcomes d'Ewing correspondent à des feuillets de petites cellules rondes, bleues et indifférenciées. Le type cellulaire à l'origine de la tumeur reste encore aujourd'hui inconnu. James Ewing avait d'abord suggéré une cellule d'origine endothéliale dans la mesure où la tumeur présentait des similarités avec des tumeurs angio-endothéliales [79]. Depuis d'autres origines ont été proposées : cellules hématopoïétiques [82], fibroblastiques [83] ou cellules souches mésenchymateuses [84]. Plus récemment l'hypothèse d'une cellule épithéliale a refait surface [85] sans permettre néanmoins de trancher définitivement sur la cellule à l'origine des sarcomes d'Ewing.

Les sarcomes d'Ewing présentent une caractéristique remarquable : le processus d'oncogénèse est initié par une seule modification génétique : une translocation chromosomale notée (11; 22) (q24 ; q12). Cette translocation chromosomale est responsable de l'existence d'un gène de fusion spécifique qui implique le plus souvent le gène EWS du chromosome 22 et le gène FLI-1 de la famille ETS localisé sur le chromosome 11 [86].



**Figure 5 : Représentation schématique de l'obtention de la protéine de fusion EWS-FLI1 par translocation chromosomale (Figure d'après [87]).**

Moins fréquemment, le gène EWS est fusionné avec d'autres gènes de la famille ETS comme ERG, ETV1, E1A-F, FEV [88]. La fusion de ces gènes conduit à l'expression par les cellules tumorales d'une protéine chimérique qui contient les acides aminés de la partie N-terminale de la protéine EWS, un puissant activateur de transcription, et le domaine de liaison de l'ADN du facteur de transcription ETS. Dans le cas le plus fréquent de fusion avec le gène FLI1, la protéine obtenue est nommée EWS-FLI1. EWS-FLI1 est une oncoprotéine (protéine fabriquée par l'organisme sur la commande d'un oncogène (gène altéré)) qui agit comme régulateur de transcription et module l'expression de centaines voire de milliers de gènes [89]. La purification et la caractérisation biochimique de la protéine EWS-FLI1 ont montré que cette dernière avait une structure désordonnée [90] et que ce désordre intrinsèque de la protéine EWS-FLI1 était important pour son activité transcriptionnelle [91]. En effet, la flexibilité et la capacité de mouvement des protéines désordonnées leur permettent de s'associer et de se dissocier rapidement de complexes protéiques. Dans le cas d'EWS-FLI1, sa structure désordonnée lui permet de participer à des complexes multiprotéiques nucléaires [92].



Bien que le rôle de EWS-FLI1 comme régulateur de transcription soit plutôt bien documenté [89], assez peu d'informations sont recensées sur la manière dont l'expression de la protéine de fusion EWS-FLI1 influe sur le comportement de la cellule et au final sur le phénotype de la tumeur chez le patient.

Il a été montré par Chaturvedi *et al.* que la protéine EWS-FLI1 n'entraînait pas l'apparition des caractéristiques classiques développées par les cellules cancéreuses pour se disséminer, telles une augmentation de l'adhésion cellulaire, de la capacité de migration et d'invasion.

En effet, il apparaît qu'EWS-FLI1 induit une désorganisation très importante des assemblages de microfilaments d'actine, fibres de stress et points d'adhésions focaux [93]. La dissémination de ces cellules cancéreuses ne se feraient ainsi pas *via* la mise en place de protrusions (filapodia, invadopodia...) mais de manière passive dans la mesure où ces dernières présentent une adhésion diminuée.

Quelques études ont montré que des altérations de la structure de filaments d'actine dans ces cellules sont corrélées avec la sous-expression de certaines protéines du cytosquelette et que la ré-expression de ces protéines permet de restaurer à la fois les microfilaments d'actine mais également les points d'adhésions focaux et d'inhiber la croissance de ces cellules sans ancrage ou au contact c'est-à-dire de réverser le phénotype tumoral. C'est le cas de l'étude menée par Amsellem *et al.* qui a permis de mettre en évidence l'effet suppresseur de tumeur de la protéine zyxine [94] ou encore du travail réalisé par Song *et al.* mettant en évidence l'inhibition de la croissance en indépendante d'ancrage des cellules NIH-3T3 exprimant EWS-FLI1 par la protéine EPLIN (*Epithelial Protein Lost In Neoplasia*) [95].

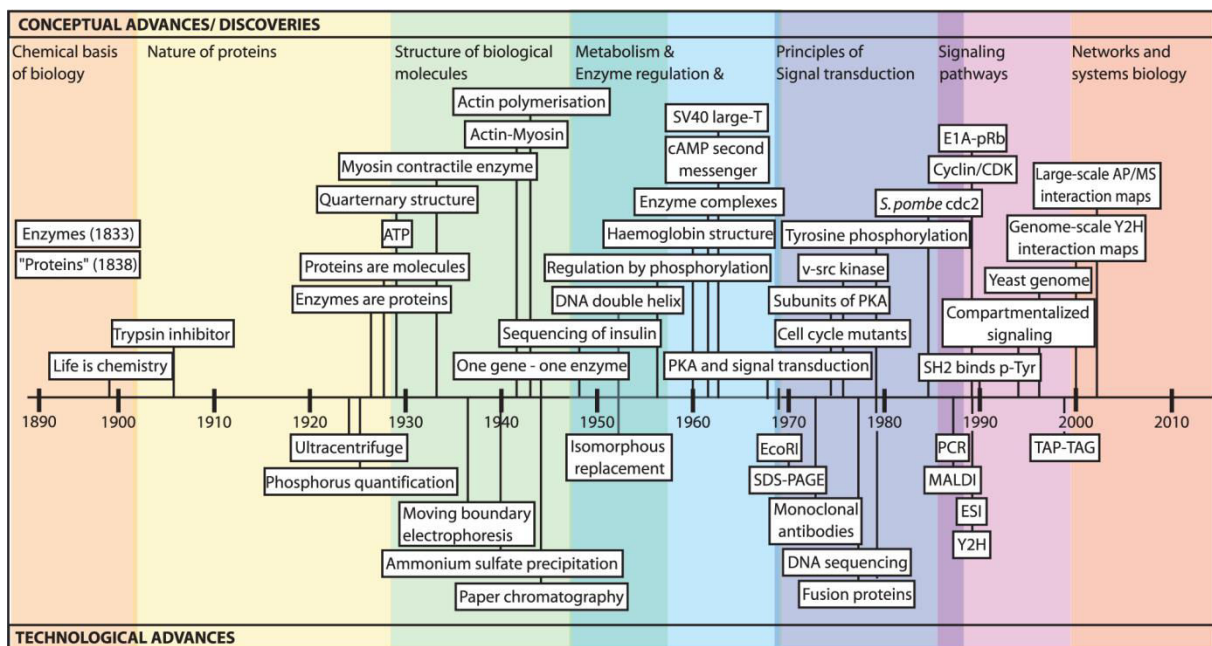
Ces études, focalisées sur le rôle d'une protéine en particulier, n'ont cependant pas permis d'établir le(s) mécanisme(s) moléculaire(s) par le(s)quel(s) EWS-FLI1 entraîne ces changements d'architecture du cytosquelette. Les voies de signalisation impliquées, aboutissant au phénotype tumoral des cellules exprimant la protéine EWS-FLI1, restent très peu connues.



## Chapitre 2 : Méthodes pour l'étude des interactions protéine-protéine : de la découverte de nouveaux partenaires à l'intégration dans les réseaux d'interactions

En raison du rôle crucial des interactions protéine-protéine en biologie (plusieurs centaines de milliers estimées chez l'homme avec entre 3 et 5 partenaires d'interactions par protéine [96], [97]), lorsque nous étudions aujourd'hui une protéine inconnue, une des premières questions que nous nous posons est : « **Quels sont ses partenaires d'interaction ?** ».

Cela n'a pas toujours été le cas. En effet l'étude des interactions protéiques a démarré lentement et a considérablement évolué au fur et à mesure des progrès techniques et conceptuels qui ont été faits dans différents domaines de la recherche en biologie. Braun P. et Gingras A-C ont retracé l'évolution des méthodes mises en œuvre pour l'étude des interactions protéine-protéine dans une revue de qualité dont est issue la Figure 6 présentée ci-dessous [96].



**Figure 6 : Frise chronologique récapitulant les évolutions de la recherche sur les interactions protéine-protéine.** La partie supérieure est consacrée aux avancées conceptuelles et aux découvertes tandis que la partie inférieure fait référence aux avancées technologiques et aux inventions (Figure d'après [96]).

L'objet de ce chapitre est de présenter les techniques qui sont aujourd'hui les plus utilisées pour étudier les interactions protéine-protéine. D'abord, nous nous intéresserons aux **méthodes ciblées de découvertes de partenaires d'interaction** et verrons comment les interactions découvertes pour une protéine d'intérêt peuvent être **validées et caractérisées**. Ensuite, nous essayerons de voir

comment des expériences de protéomique différentielle sont susceptibles de nous renseigner sur les interactions protéine-protéine. Finalement, nous verrons comment ces différentes méthodes se complètent pour venir étoffer la connaissance générale des interactions protéine-protéine et de leur organisation en réseau. En perspective, nous aborderons la nécessité de tenir compte des effets dynamiques dans les études d'interaction protéine-protéine.

Pour faciliter la lecture de ce chapitre qui pourrait apparaître un peu dense, vous trouverez ci-dessous un schéma récapitulatif (Figure 7) qui résume et classe différentes techniques d'étude des interactions protéine-protéine auxquelles il est possible de faire appel, dont un certain nombre tel que le pontage covalent, la purification d'affinité couplée à la masse et la protéomique différentielle par marquage métabolique SILAC que j'ai utilisées pour mon travail de thèse.

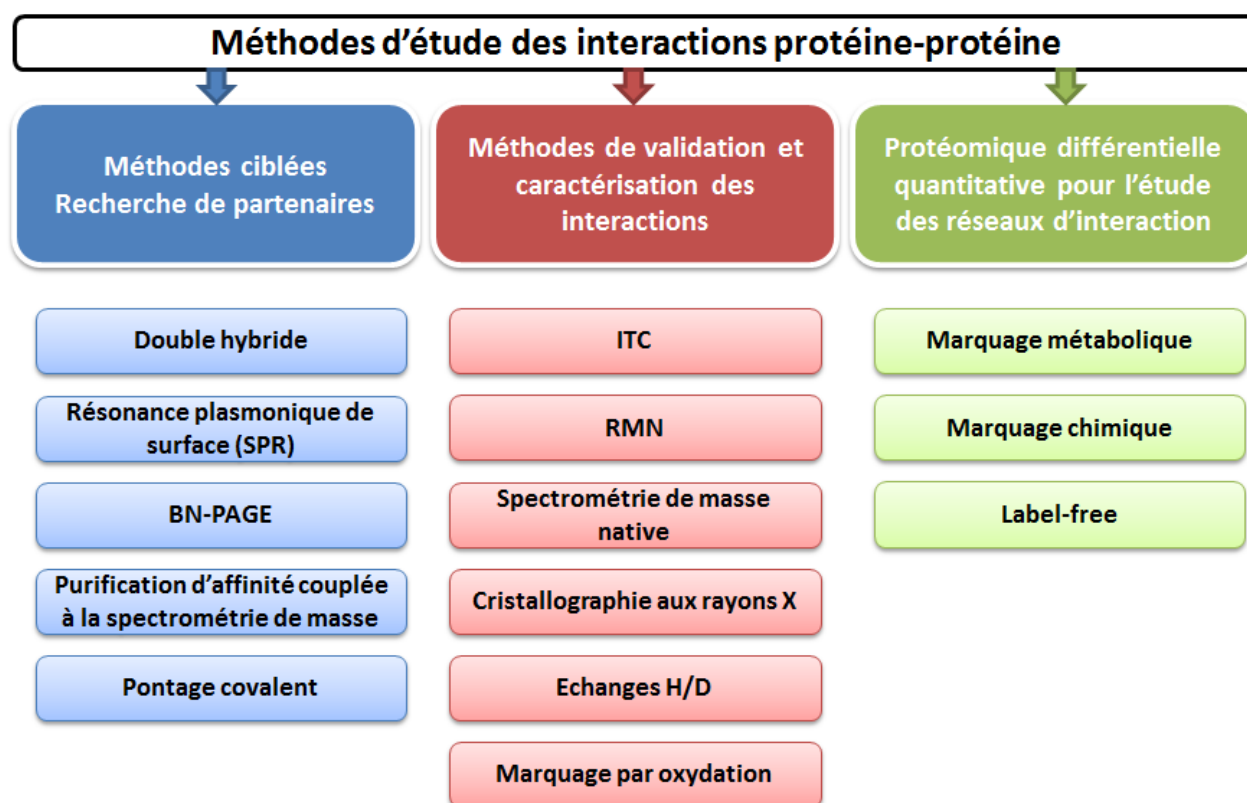


Figure 7 : Schéma récapitulatif des différentes méthodes d'étude d'interactions protéine-protéine qui seront abordées dans ce chapitre.

## I. Méthodes ciblées de découverte de partenaires d'interaction

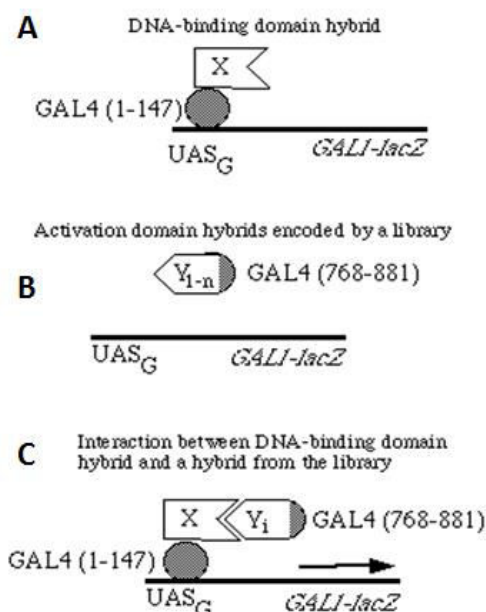
### 1. Systèmes cellulaires d'expression de double hybride

Le système « double hybride » (*two-hybrid system*) décrit pour la première fois en 1989 par Field et Song dans la levure *saccharomyces cerevisiae* [98] est une technique de biologie moléculaire permettant de détecter une interaction entre deux protéines dans un milieu intracellulaire. On estime que 50 % des interactions entre protéines décrites dans la littérature ont été détectées grâce au système « double hybride ».

#### a. Principe de la méthode originale des doubles hybrides chez la levure

Le concept général repose sur le fait que la plupart des facteurs transcriptionnels activateurs interagissant avec l'ADN sont composés d'au moins deux domaines ayant des fonctions spécifiques : un domaine de liaison à l'ADN (DBD ou *DNA-binding domain*) et un domaine d'activation (AD ou *activation domain*), connectés l'un à l'autre par une région charnière, l'ensemble donnant une architecture tripartite de type « DBD-région charnière-AD » douée d'une fonction de transactivation.

Chacun des domaines pris séparément est inefficace. Ils doivent être associés pour activer la transcription soit naturellement grâce à la région charnière, soit artificiellement grâce à des intermédiaires de nature diverse pouvant être notamment des interactions protéine-protéine. En effet, il a été démontré expérimentalement que deux protéines chimériques, l'une comportant un domaine DBD relié à un polypeptide X (hybride DBD-X) et l'autre comportant un domaine AD relié à un polypeptide Y (hybride AD-Y) pouvaient reconstituer un facteur transcriptionnel actif si les régions polypeptidiques X et Y sont capables d'interagir l'une avec l'autre. Ces deux protéines chimériques permettant de tester l'interaction protéine-protéine sont exprimées dans la levure par transfection des constructions plasmidiques appropriées. L'exemple de la reconstitution d'une protéine GAL4 fonctionnelle (= comportant à la fois un DBD et un AD) capable d'activer la transcription du gène *GAL1-lacZ* codant pour l'enzyme galactosidase est présenté Figure 8.



**Figure 8 : Exemple de stratégie utilisée pour détecter une interaction protéine-protéine avec un système double hybride rétablissant l'activité de la protéine GAL4. (Figure d'après[99])**

A = La partie N-terminale de la protéine GAL4 (1-147) lui permet de se lier à la séquence d'activation du gène (UAGs *upstream activation sequence*) *Gal1-lacZ*. Cette partie N-terminale de la protéine a été reliée à un polypeptide X.

B = La partie C-terminale de la protéine GAL4 (768-881) qui correspond à son domaine d'activation est reliée à un polypeptide Y.

C = Si les polypeptides X et Y interagissent, leur interaction va permettre de rétablir le lien entre les deux domaines de la protéine GAL4 et ainsi rétablir son activité de transcription du gène *Gal1-lacZ*.

D'autres systèmes reposent sur la complémentarité de deux domaines d'une protéine. Par exemple, des enzymes ou des protéines fluorescentes sont coupées en deux domaines et lorsque les protéines auxquelles ces domaines sont fusionnés interagissent l'une avec l'autre, l'activité enzymatique (comme avec la  $\beta$ -galactosidase) ou la fluorescence (avec la protéine fluorescente jaune) sont ré-observées. Ces expériences basées sur la fluorescence par complémentarité bimoléculaire sont de plus en plus utilisées pour détecter des interactions protéine-protéine dans le contexte de cellules entières [100], [101].

### **b. Limites du système double hybride**

La sélection des clones positifs reposant sur l'activation de la transcription d'un gène rapporteur, cela implique que l'interaction entre les deux protéines de fusion se produise dans le noyau de la cellule. Cette caractéristique s'accompagne de plusieurs conséquences notables qu'il convient de prendre en compte. Le milieu au sein duquel s'établit l'interaction protéine-protéine (nucléoplasme) peut être différent de celui où résident naturellement les protéines (cytoplasme, milieu extracellulaire, membrane, compartiment subcellulaire...) ce qui peut d'emblée influencer sur le fait que les protéines vont entrer ou non en contact (le pH, l'état redox, le type d'ions et leur concentration ainsi que d'autres facteurs jouent un rôle important dans les interactions). L'environnement nucléoplasmique des protéines qui ne sont pas forcément nucléaires peut atténuer, ou, au contraire, favoriser les interactions entre protéines. Enfin pour interagir avec un partenaire, certaines protéines doivent subir des modifications post-traductionnelles (phosphorylation, acétylation...) nécessitant des enzymes spécifiques d'un type cellulaire donné ou ayant une localisation particulière.

Ainsi parmi toutes les protéines candidates à une exploration sur le principe du double hybride chez la levure, certaines classes de protéines se prêtent mal ou très mal à l'approche classique, les protéines membranaires surtout mais aussi les protéines fibrillaires et certains facteurs transcriptionnels (de par leur aptitude potentielle à activer la transcription).

Des méthodes alternatives peuvent être utilisées. Elles jouent sur le choix du type cellulaire ou sur le site subcellulaire où se déroule l'interaction.

### **c. Cellules hôtes**

Le test original a été développé avec des levures mais théoriquement toutes les cellules sont susceptibles de servir de tube à essai pour la détermination d'interactions protéine-protéine. Les cellules mammifères communément utilisées appartiennent généralement à des lignées telles que CHO, COS [102], HEK-293, Hela, NIH 3T3.

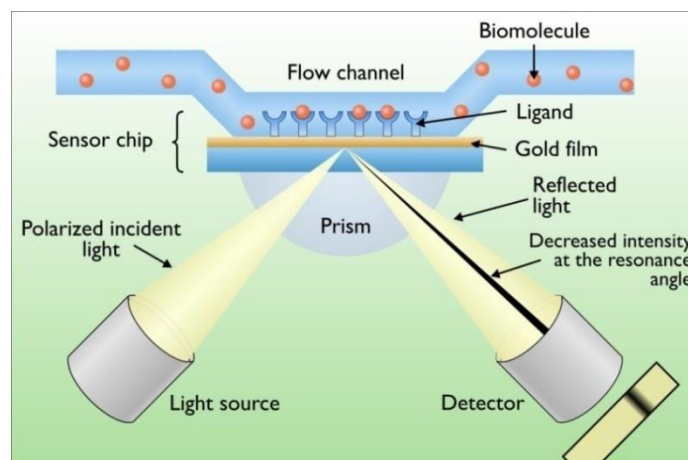
## **2. Résonance plasmonique de surface (*Surface plasmon resonance, SPR*)**

La technique de SPR a été pour la première fois utilisée comme biocapteur par Leidberg *et al.* en 1983 [103]. Cette méthode permet de détecter la formation de complexes en suivant des changements d'angle d'incidence ou d'intensité d'un faisceau lumineux monochromatique lors d'une réflexion totale. Les partenaires d'une protéine donnée en solution interagissent avec la protéine

d'intérêt fixée de façon covalente ou non sur un film organique d'une épaisseur de 100 nm, lui-même lié à une surface métallique de faible épaisseur.

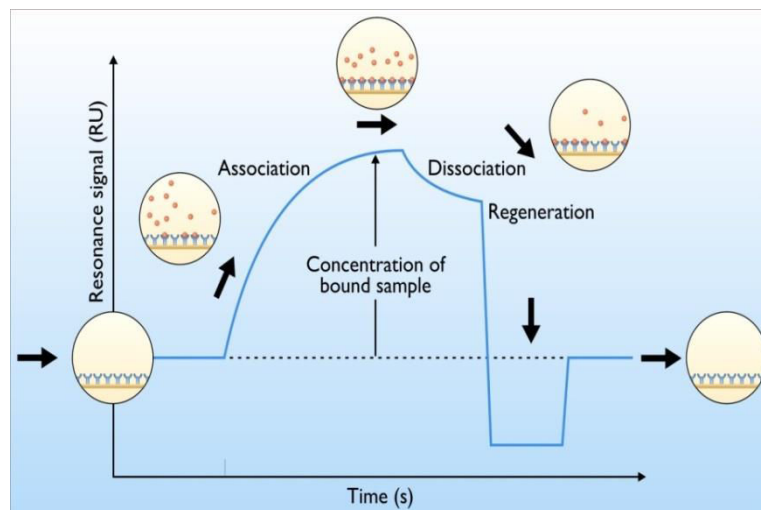
Les phénomènes physiques impliqués dans la détection de partenaires sont complexes. On peut les détailler comme suit : La réflexion totale d'un faisceau laser à l'interface verre/métal génère une oscillation de densité d'électrons quasi-libres (plasmon) du métal. Cette oscillation de densité génère à son tour une onde électromagnétique évanescente se propageant sur une épaisseur de ~100 nm à partir de l'interface. Toute modification de composition dans cette épaisseur (protéine avec et sans partenaires) entraîne une modification de la fréquence de la résonance plasmonique et donc une modification de la réflexion initiale (angle et/ou intensité).

Une expérience de SPR se déroule ainsi classiquement de la manière suivante : un ligand d'intérêt (un peptide ou une protéine par exemple) est immobilisé covalamment sur un polymère de dextran lui-même fixé covalamment sur une surface en or et exposé au flux d'une solution de protéines contenant des partenaires d'interaction potentiels. Les protéines qui interagissent avec ce ligand sont retenues à la surface du polymère et vont induire une modification de l'angle de résonance de la lumière en changeant l'indice de réfraction au fur et à mesure de leur accumulation par interaction avec le ligand d'intérêt. Cette modification de l'angle de résonance se traduira par une diminution de l'intensité du signal pour l'angle de résonance initial. Un schéma de principe de la technique est présenté Figure 9.



**Figure 9 : Schéma de principe de la résonance plasmonique de surface (SPR)** (d'après le site <http://elte.prompt.hu>)

Il existe une relation linéaire entre la diminution de l'intensité pour l'angle de résonance et la concentration en protéine à la surface. Ainsi il est possible de déterminer la constante d'association  $k_a$  en suivant l'évolution du signal de SPR au cours du temps comme schématisé Figure 10. La constante de dissociation  $k_d$  peut également être mesurée en lavant la surface avec le tampon seul. Ensuite, la constante de formation du complexe peut être obtenue  $K = k_a/k_d$ .



**Figure 10 : Représentation schématisée d'une mesure de SPR** (d'après le site <http://elte.prompt.hu>)

Les articles de Malmqvist [104] et de Jonsson et al. [105] constituent de bonnes références pour ceux qui souhaitent avoir une explication plus détaillée sur l'utilisation de la SPR pour l'étude d'interaction bimoléculaires tandis que l'article de Karlsson [106] reprend les différentes équations traduisant ces phénomènes.

Un dispositif basé sur la SPR pour l'étude d'interaction de biomolécules est commercialisé depuis 1984 sous le nom de Biacore. Cet appareil présente plusieurs avantages [13]: (1) il nécessite peu de matériel biologique. Typiquement seulement 1 à 10  $\mu\text{g}$  de protéine doivent être immobilisé sur la puce qui peut ensuite être réutilisée jusqu'à 50 fois après élimination des protéines adhérentes. (2) La méthode est très rapide, de l'ordre de 10 min pour une expérience. (3) Cette technique ne nécessite aucune modification préalable des protéines c'est-à-dire aucun marquage par des étiquettes d'affinité ou fluorescentes. (4) Des interactions peuvent être mises en évidence même dans des mélanges complexes de protéines. (5) A la fois les constantes d'association et de dissociation peuvent être obtenues rapidement. (6) Cette méthode est applicable à une large gamme de concentrations en protéines.

La SPR a été utilisée avec succès pour suivre des interactions peptide-protéine et a permis, par exemple, de déterminer l'interaction de différents domaines SH2 avec deux peptides dérivés du facteur de croissance des plaquettes (PDGF) porteurs de phosphorylations de leur acides aminés tyrosine [107].

Cette technologie peut être utilisée pour mieux caractériser des interactions protéine-protéine déjà connue en déterminant la constante de formation de ce complexe protéique mais peut également permettre d'isoler de nouveaux partenaires d'interaction qui, une fois élués et digérés, seront identifiés par spectrométrie de masse en tandem. Kikuchi *et al.* [108] ont montré la faisabilité de ce type d'approche dite de « ligand fishing » pour l'identification de nouveaux partenaires de la protéine p53. La SPR fait ainsi également partie des méthodes pouvant conduire à la découverte de nouvelles interactions protéines-protéines.



Le couplage avec la spectrométrie de masse (MS) peut être indirect (élution des partenaires) ou direct (analyse en MALDI de la puce biocapteur). Compte tenu des écarts de sensibilités des deux techniques (~1 ng en SPR et 10 ng en MS) leur couplage reste aujourd'hui encore délicat en dehors des interactions anticorps-antigène.

### 3. Blue Native PAGE

La technique *blue native polyacrylamide gel electrophoresis* (BN-PAGE) mérite d'être mentionnée parmi les méthodes de découvertes d'interactions protéine-protéine. La technique Blue Native PAGE (BN-PAGE) mise au point par Schägger *et al.* [109] reprend le principe du gel SDS-PAGE qui permet de séparer les protéines en fonction de leur taille. Au lieu des molécules de SDS qui se fixent aux protéines et leur confèrent une charge négative en les dénaturant, c'est le bleu de Coomassie G qui va permettre, dans le cas de la technique BN-PAGE, de charger négativement les protéines mais sans les dénaturer. Il est également possible avec cette technique de maintenir des interactions non-covalentes et donc d'isoler des complexes protéiques. Cette technique a ainsi été utilisée pour établir des interactions protéine-protéine. Elle s'est, par exemple, révélée très utile pour étudier les complexes de respiration des mitochondries chez les plantes ou chez la levure [110], [111].

A l'issue de la séparation sur BN-PAGE qui permet de déterminer la masse des protéines dans leur état natif, de définir les états d'oligomérisation et d'identifier des interactions protéine-protéine, il est possible de réaliser (1) une extraction des protéines par électroélution et une immunodétection après électrotransfert, (2) une migration dans une seconde dimension en SDS-PAGE ou même (3) une migration dans une seconde dimension en électrofocalisation (IEF) suivie d'une séparation par SDS-PAGE dans une troisième dimension. Ainsi les sous-unités des complexes isolés en BN-PAGE peuvent être dissociées et pourront être identifiées par la suite par spectrométrie de masse par exemple [112].

### 4. Méthodes de purification d'affinité couplées à la spectrométrie de masse (AP-MS)

Les méthodes de découvertes de partenaires d'interaction basées sur la purification par affinité consistent dans un premier temps à isoler une **protéine d'intérêt (l'appât)** ainsi que ses **partenaires d'interaction (les proies)** grâce à une molécule immobilisée sur un support solide (une résine par ex.) qui va permettre de les capturer. Dans un deuxième temps les partenaires d'interaction isolés vont être identifiés par spectrométrie de masse.

Ainsi ces approches de purification d'affinité couplées à la spectrométrie de masse sont généralement désignées par l'acronyme AP-MS (*Affinity Purification coupled with Mass Spectrometry*).

Nous allons tout d'abord nous focaliser sur les différentes méthodes de purification d'affinité avant d'aborder l'identification des partenaires capturés par spectrométrie de masse.

Au cours des dernières années, les méthodes de purification d'affinité ont revêtu plusieurs formes en raison du développement de nombreux protocoles adaptés à différents contextes biologiques.

## Introduction : Méthodes d'étude des interactions protéine-protéine

On distingue deux grandes familles de protocoles en fonction de l'ordre des événements :

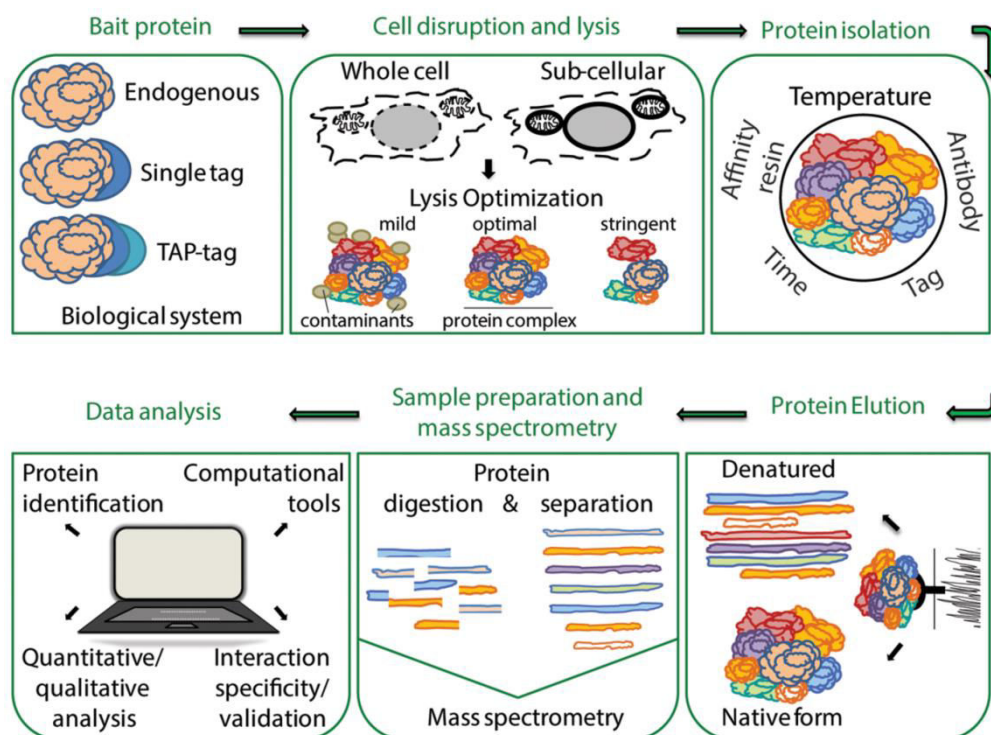
- Soit la protéine d'intérêt appât est purifiée sur le support d'affinité en même temps que ses partenaires d'interaction : on parle d'**immunoprécipitation**.
- Soit la protéine d'intérêt est d'abord immobilisée sur le support d'affinité et ensuite mise en présence d'un mélange complexe de protéines contenant des partenaires potentiels : on parle alors d'un protocole de **pull-down**.

### a. Purifications d'affinité

#### Expériences d'immunoprécipitation

Il y a plusieurs étapes universelles dans une expérience d'immunoprécipitation qui sont listées ci-dessous et illustrées par la Figure 11.

- (1) L'obtention, la préparation de cellules ou de tissus contenant la protéine d'intérêt sous forme endogène ou porteuse d'une étiquette.
- (2) La lyse des cellules ou du tissu et la solubilisation des complexes protéiques grâce à des tampons de lyse adaptés (conditions non dénaturantes pour préserver les complexes).
- (3) L'isolement de la protéine d'intérêt et de ses partenaires d'interaction par purification d'affinité.
- (4) L'élution des complexes protéiques purifiés.
- (5) L'analyse des protéines isolées par spectrométrie de masse pour l'identification/caractérisation et la quantification des interactions protéine-protéine.



**Figure 11 : Description schématique des étapes universelles d'une expérience d'immunoprécipitation couplée à la spectrométrie de masse [113]:**

## Introduction : Méthodes d'étude des interactions protéine-protéine

Les protocoles d'immunoprécipitation varient en fonction de la protéine appât utilisée au départ qui peut être soit endogène, soit porter une étiquette (*epitope tag*), mais qui peut également avoir été modifiée par deux étiquettes distinctes comme pour les expériences de TAP-tag.

### ➤ Purification d'affinité avec une protéine endogène

L'immunoprécipitation avec une **protéine endogène** repose sur la capacité d'isoler la protéine appât sur un support par l'intermédiaire d'un anticorps spécifique de cette protéine. L'avantage de cette méthode est la possibilité de capturer la protéine dans son état, avec une abondance physiologique et en interaction avec ses partenaires pour une multitude de systèmes (cellules, tissus, modèles animaux) sans avoir besoin de réaliser de clones ou de marquage.

Les inconvénients sont la disponibilité, la spécificité et l'affinité des anticorps utilisés pour reconnaître la protéine d'intérêt. La réactivité croisée et le coût de ces anticorps constituent souvent un frein à leur utilisation.

### ➤ Purification d'affinité avec une protéine porteuse d'une étiquette (*epitope tag*)

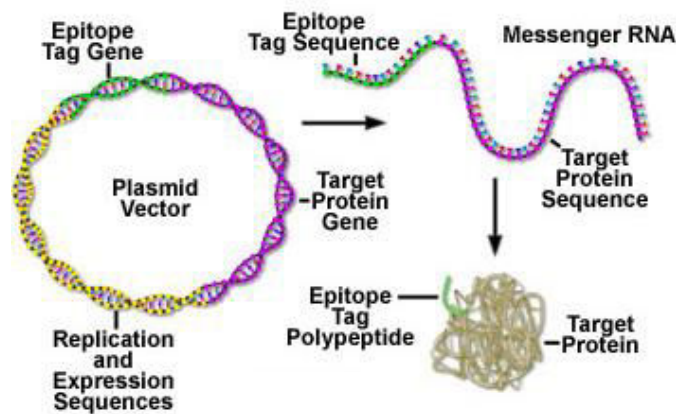
En alternative à l'utilisation de la forme endogène de la protéine, il est fréquent de faire appel à une étiquette (*epitope tag*). Ces étiquettes sont de courtes séquences peptidiques qui ont une forte affinité avec des anticorps. Elles sont généralement dérivées de gènes viraux ce qui explique leur forte immunoréactivité. De nombreuses étiquettes ont déjà été utilisées pour les expériences d'immunoprécipitation [114], [115]. Parmi elles, se trouvent des peptides étiquettes tels que le FLAG, 6 histidines (His)<sub>6</sub>, c-Myc et l'hémagglutinine ou des étiquettes de tailles plus importantes telles que la protéine fluorescente verte (*green fluorescent protein* GFP) ou la glutathione-S-transférase (GST). Les étiquettes les plus communes ainsi que leurs caractéristiques sont rassemblées dans le Tableau 1.

**Tableau 1 : Etiquettes communément utilisées en purification d'affinité (Figure d'après [114]).**

Name	Sequence	Detection	Purification
FLAG	DYKDDDDK	M1, M2, M5	Immunoaffinity
6 × His	HHHHHH	Anti-His	Metal affinity
HA	YPYDVPDYA	12CA5	Immunoaffinity
c-myc	EQKLISEEDL	9E10	Immunoaffinity
GST	220 aa GST	Anti-GST	Glutathione
Protein A	IgG-binding domain	IgG	IgG
CD	18 aa exon	12CA5	Immunoaffinity
Strep-tag	WSAPQFEK	Strep-Tactin	Strep-Tactin
MBP	Maltose-binding protein	Anti-MBP	Maltose
CBD	Chitin-binding domain	Anti-CBD	Chitin
S-tag	S-peptide	Anti-S peptide	S-peptide
Avitag	GLNDIFEAQKIEWHE	Avidin	Avidin
CBP	CBP peptide	Anti-CBP	Calmodulin
TAP	Calmodulin- and IgG-binding domains	Anti-CBP	Calmodulin and IgG
SF-TAP	Strep Tag II and FLAG	Anti-FLAG	Strep-Tactin

GST, glutathione-S-transferase; CBP, calmodulin-binding peptide.

L'étiquette est fusionnée à la protéine en insérant le gène de cette protéine dans le vecteur contenant le gène codant pour la séquence de l'étiquette comme illustré dans la Figure 12.



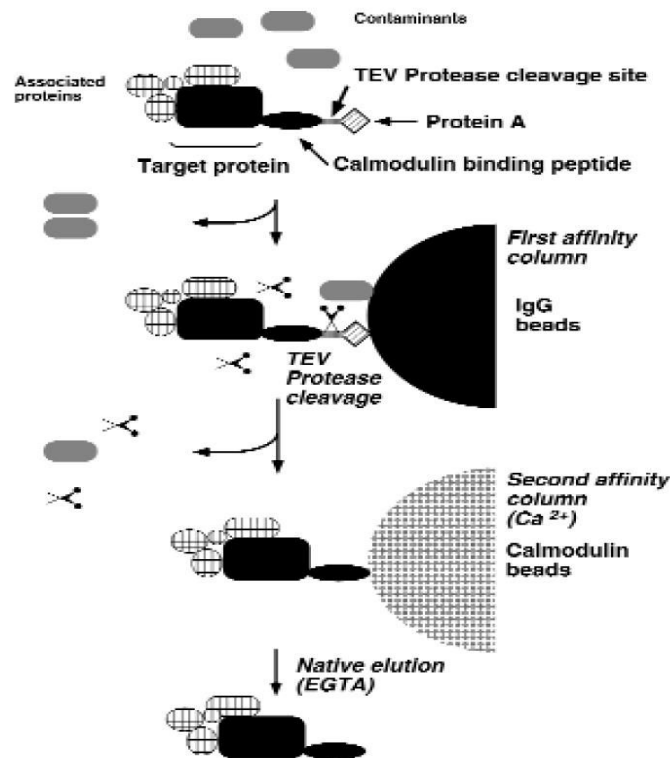
**Figure 12 : Obtention d'une protéine recombinante porteuse d'une étiquette** (site olympusmicro.com rubrique epitope tagging)

La méthode de purification d'affinité avec une protéine porteuse d'une étiquette est indépendante des propriétés de liaison de la protéine et peut donc être appliquée à la fois à des protéines de forte ou de faible abondance. Cette approche est généralement plus robuste car les conditions expérimentales sont optimisées pour une étiquette donnée qui peut être ensuite utilisée pour étudier différentes protéines. Un inconvénient de cette méthode est le temps nécessaire à l'obtention de la protéine de fusion qui doit être générée et introduite dans le système biologique d'intérêt. De plus, il est très important de valider la protéine « taggée » de manière à confirmer que l'étiquette n'interfère ni avec la fonction de la protéine endogène, ni avec sa localisation ou ses propriétés [116].

➤ **Méthode de purification en tandem TAP-tag (double tag)**

Tandis que les méthodes de purification d'affinité comportant une seule étape ont fait leurs preuves pour la capture d'interactions protéiques de faible intensité et de faible abondance, les techniques de purification d'affinité en tandem ont été initialement introduites pour préserver les interactions stables tout en réduisant les interactions non spécifiques [117].

La première expérience de TAP-tag schématisée Figure 13 était basée sur l'utilisation d'une protéine porteuse d'une double étiquette : un peptide liant la calmoduline (*Calmoduline Binding Peptide* CBP) suivi d'une unité IgG de la protéine A de *Staphylococcus aureus* (PrA) et reliés par une séquence peptidique correspondant au site de clivage par la protéase TEV. La protéine de fusion correspondante et ses partenaires d'interaction sont purifiés à partir d'extraits cellulaires par une première sélection sur une matrice IgG. Après différents lavages, la protéase TEV est ajoutée pour cliver la protéine de fusion de son support. L'éluat est ensuite incubé avec des billes greffées par de la calmoduline en présence de calcium. Cette seconde étape de purification d'affinité permet d'éliminer la protéase TEV ainsi que les traces de contaminants restantes après la première purification. A l'issue de nouveaux lavages, la protéine de fusion et ses partenaires stables sont élués avec de l'EGTA (*ethylene glycol tetraacetic acid*) qui va chélater les ions calcium et induire un changement de conformation de la calmoduline permettant la libération de la protéine de fusion.



**Figure 13 : Schéma de l'expérience de TAP-tag initiale (Figure tirée de [117]).**

Bien que faisant référence au départ à la combinaison spécifique des étiquettes peptidiques de la protéine A et de la calmoduline développée par l'équipe de Séraphin, le terme de purification TAP-tag est aujourd'hui utilisé pour décrire toutes les techniques qui font appel à deux étiquettes et à deux étapes de purification d'affinité successives. D'excellents articles de revue décrivent les différentes stratégies employées [118], [119]. Il est important de garder à l'esprit que la taille relativement importante de ces tags et de leurs connexions aux régions N- ou C-terminales peuvent compliquer fortement l'interprétation et la filiation des nombreux partenaires. Il peut ainsi être utile de comparer les résultats obtenus en plaçant ces tags soit en N-, soit en C-ter de la protéine.

### **Expériences de pull-down**

Dans une expérience de pull-down, la protéine appât qui porte une étiquette est immobilisée sur un support par le biais de cette étiquette. Un support « d'affinité secondaire » est ainsi généré pour la purification de protéines partenaires de la protéine appât. Ce support sur lequel est immobilisée la protéine appât est incubé avec un milieu contenant un mélange complexe de protéines partenaires potentielles tel qu'un lysat cellulaire. Cette méthode est plus simple à mettre en œuvre car elle ne nécessite pas forcément d'exprimer une protéine de fusion porteuse d'une étiquette. En effet, l'étiquette peut être ajoutée à la protéine après sa purification. En revanche, il se peut que la fixation préalable de la protéine appât entraîne une contrainte conformationnelle de celle-ci conduisant au masquage d'un ou plusieurs sites d'interaction de cette protéine avec son ou ses partenaires de liaison. Cette méthode est également applicable à des petites molécules de synthèse, peptides ou autres, dans lesquelles une étiquette telle qu'une biotine peut avoir été introduite. L'utilisation de ce type de stratégie destinée à comprendre l'effet d'une petite molécule en se basant sur

l'identification de ses partenaires d'interaction porte le nom de **protéomique chimique** (*chemical proteomics*) [120].

### b. Couplage à la spectrométrie de masse haut débit

L'utilisation de ces stratégies de purification d'affinité à grande échelle pour la découverte de nouveaux partenaires d'interaction n'aurait pas pu se développer sans les progrès importants réalisés en spectrométrie de masse. A la fin des années 1990, la spectrométrie de masse a su surmonter différentes contraintes technologiques liées à l'ionisation (ESI, MALDI), à l'analyse (TOF, puis orbitrap) et au couplage avec les techniques de séparation (LC) qui lui ont permis de devenir la méthode de choix pour l'identification rapide de protéines digérées issues de morceaux de gels ou sous forme de mélanges complexes en solution.

La protéomique haut-débit dite « shotgun » est aujourd'hui la technologie la plus puissante pour l'analyse de mélanges complexes de protéines. Elle fait appel à une **stratégie dite « bottom-up »** reposant sur la digestion enzymatique d'ensembles complexes de protéines suivie de la séparation en chromatographie liquide haute performance des peptides tryptiques obtenus qui sont ensuite analysés par spectrométrie de masse en tandem. La spectrométrie de masse en tandem comprend trois étapes. Au cours de la première étape, les rapports  $m/z$  des peptides sont mesurés, dans une seconde étape les peptides sélectionnés sont fragmentés et enfin dans une troisième, les rapports  $m/z$  des différents fragments d'un peptide sont mesurés. Des séparations chromatographiques successives peuvent être réalisées grâce à l'utilisation de phases stationnaires de différentes natures (par exemple échangeuse de cations et phase inverse) directement en ligne avec l'analyse par spectrométrie de masse. Cette technique est communément appelée MudPIT pour *Multi Dimensional Protein Identification Technology* [121]. L'identification des protéines est ensuite généralement réalisée en comparant les spectres de fragmentation expérimentaux avec ceux générés *in silico* et en sélectionnant ainsi la séquence peptidique la plus susceptible de correspondre. Cette recherche se fait *via* des moteurs de recherche tels que SEQUEST [122] ou MASCOT [123]. Les identifications sont ensuite filtrées en fonction de scores relatifs à leur qualité et en estimant un taux de faux positif (FDR *False Discovery Rate*).

Depuis quelques années une autre stratégie gagne rapidement du terrain en protéomique : il s'agit de la **stratégie dite « top-down »**. Basée sur l'analyse de protéines entières, elle permet de s'affranchir d'un certain nombre de défauts de l'analyse bottom-up. En effet, en analyse bottom-up une première difficulté réside dans le fait que certains peptides ne sont pas spécifiques d'une protéine et que, par conséquent, l'identification de ces peptides ne permet pas de trancher sur l'identité de la protéine. D'autre part, avec la stratégie bottom-up, la couverture de séquence est bien souvent incomplète et nous fait passer à côté d'informations potentiellement intéressantes telles que l'intégrité d'une protéine et l'existence de modifications post-traductionnelles. Finalement après digestion enzymatique l'information sur la relation entre les différents peptides d'une protéine est perdue : il ne sera, par exemple, pas possible de dire si deux peptides phosphorylés attribués à une même protéine peuvent l'être simultanément ou pas.

L'essor de la stratégie top-down en protéomique correspond au développement récent de méthodes de séparation efficaces des protéines (électrophorèse en tube de gel, focalisation isoélectrique, nouvelles phases stationnaires en chromatographie), de spectromètres de masse à la fois très



résolutifs, avec une excellente précision en masse et une bonne sensibilité (FT-ICR, LTQ-Orbitrap) [124]. En top-down, l'ionisation des protéines entières est majoritairement réalisée en électrospray. Ces grosses molécules sous formes très multichargées peuvent ensuite être fragmentées par différentes méthodes tels que le CID (*collision induced dissociation*), l'ECD (*electron capture dissociation*) ou encore le NSD (*nozzle skimmer dissociation*). La stratégie top-down est particulièrement bien adaptée à la caractérisation complète des protéoformes, c'est-à-dire des différentes formes moléculaires d'une protéine résultant de la combinaison de différentes variations génétiques, d'épissages alternatifs et de modifications post-traductionnelles [125]. Aujourd'hui, en réalisant des étapes de préfractionnement il est possible d'identifier plus de 1000 protéines d'un lysat cellulaire [126]. D'ici quelques années cette stratégie pourrait véritablement constituer une alternative viable à la stratégie bottom-up notamment pour les études d'interaction protéine-protéine faisant appel à la purification d'affinité.

L'avantage des techniques de purification d'affinité est leur capacité à mettre en évidence des interactions multidirectionnelles et pas seulement des interactions binaires comme avec le système double hybride. Cependant la difficulté essentielle de ces méthodes réside dans l'interprétation des données obtenues. En effet, il faut d'une part, être en mesure d'évaluer quelles interactions sont de « vraies » interactions et d'autre part, être en mesure de replacer les informations obtenues dans le contexte cellulaire pour comprendre la signification biologique du résultat.

### c. Détermination de la spécificité d'une interaction

Les protéines peuvent être impliquées dans de nombreuses interactions directes ou indirectes, régulées à la fois dans l'espace et dans le temps et critiques pour la réalisation de leur fonction. Au regard de la complexité des interactions, c'est un défi d'être en mesure de distinguer les interactions spécifiques des nombreuses associations non-spécifiques potentielles. Ainsi des efforts importants ont récemment été faits pour développer des approches permettant d'évaluer la spécificité des interactions protéiques.

Les **interactions non-spécifiques** sont de deux types : celles qui sont liées aux interactions avec la résine du support, avec l'anticorps ou l'étiquette utilisée pour la purification d'affinité d'une part, et celles qui résultent des associations non-spécifiques (physico-chimiques sans rôle biologique) avec la protéine d'intérêt et qui sont par conséquent beaucoup plus difficiles à éviter.

Tandis que le premier type d'interactions non spécifiques est lié à la nature du support d'affinité et peut être limité en adaptant le choix du support, le second type d'interactions non-spécifiques naît de la modification du contexte moléculaire lorsqu'on passe de l'*in vivo* à l'*in vitro*. En effet les interactions non-spécifiques avec la protéine d'intérêt dans une expérience de purification par affinité sont en partie liées au fait que les protéines à l'issue de la lyse cellulaire ne conservent pas la localisation cellulaire qu'elles avaient *in vivo* et obtiennent ainsi des opportunités de liaisons qu'elles n'avaient pas auparavant.

De plus, une fois co-isolés sur un support, la protéine d'intérêt et ses partenaires d'interactions peuvent présenter des sites d'interactions pour des protéines du lysat dites « collantes », abondantes ou capables de reconnaître des domaines d'interactions.

Il est possible d'influer sur la présence de ces interactions non-spécifiques de différentes sources en modifiant différents paramètres expérimentaux. **Le premier élément critique à optimiser est la lyse cellulaire.** Dans le cas d'une expérience d'immunoprécipitation, la lyse doit permettre de libérer en solution la protéine d'intérêt en complexe avec ses partenaires sans perturber ces interactions et sans dénaturer les protéines. Les tampons de lyse doivent être choisis de manière à maintenir les interactions fortes ou faibles tout en limitant les interactions non-spécifiques. Ainsi des tampons de lyse dits stringents c'est-à-dire avec des concentrations fortes en sel et/ou en détergent vont permettre de conserver majoritairement les interactions fortes tandis que des tampons plus doux peuvent être élaborés pour conserver des interactions plus faibles. L'inconvénient majeur des **tampons de lyse stringents**, parfois nécessaires pour avoir accès à certaines protéines des membranes ou vésicules, est qu'ils peuvent conduire à la **dénaturation des protéines** qui va être à l'origine d'un certain nombre d'association non-spécifiques, comme par exemple avec des protéines de choc thermique (HSP pour *heat shock proteins*) qui ont la caractéristique de se lier aux protéines dénaturées [127]. Les détergents utilisés en biochimie appartiennent à trois catégories : les détergents non-ioniques, zwitterioniques ou ioniques. Différents exemples de ces détergents sont présentés dans le tableau ci-après (Tableau 2).

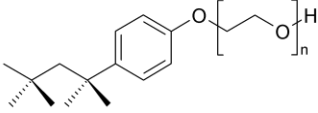
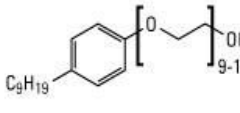
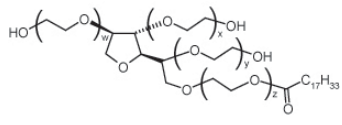
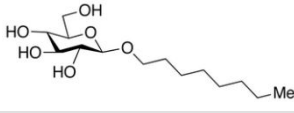
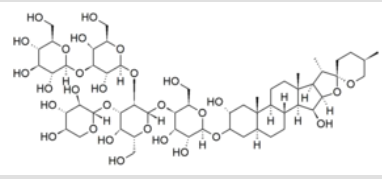
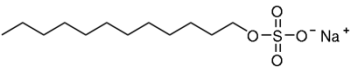
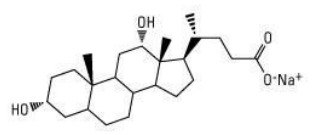
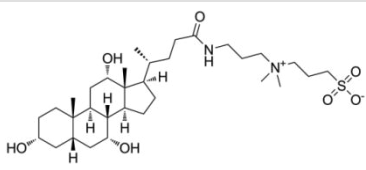
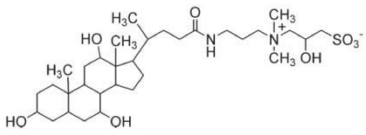
Les **détergents non-ioniques** tels que le Triton X-100 ou le NP-40 ont des têtes hydrophiles non chargées. Ils sont considérés comme des détergents doux dans la mesure où ils cassent les associations protéine-lipide ou lipide-lipide mais pas les interactions protéine-protéine et pour la plupart ne dénaturent pas les protéines. Ainsi, avec les détergents non-ioniques les protéines sont solubilisées, isolées dans leur état natif et actif en association avec leurs partenaires d'interaction. Ces détergents non-ioniques sont ainsi les plus utilisés pour isoler des protéines membranaires.

Les **détergents zwitterioniques** tels que le CHAPS et CHAPSO présentent eux une tête hydrophile avec à la fois des charges positives et négatives en nombre égal leur conférant au final une charge nulle. Ils sont plus stringents que les détergents non-ioniques et présentent l'avantage d'avoir une concentration micellaire critique (CMC) élevée ce qui permet de les éliminer efficacement par dialyse.

Les **détergents ioniques**, tels que le SDS, présentent eux une tête chargée positivement ou négativement. Ce sont des détergents très stringents qui vont casser les interactions protéine-protéine. Ils sont très efficaces pour lyser les différents compartiments cellulaires et solubiliser la majeure partie des constituants de la cellule. Ils sont cependant souvent relativement difficiles à éliminer et peuvent gêner voire complètement empêcher la digestion enzymatique des échantillons.



## Introduction : Méthodes d'étude des interactions protéine-protéine

Détergent	Type	Exemple de structure	MW mono (micelle)	CMC mM (%w/v)	Dialysable	Compatibilité digestion trypsine
Triton X-100	Non-ionique		647 (90K)	0.24 (0.0155)	Non	1%
NP-40	Non-ionique		617 (90K)	0.29 (0.0179)	Non	1%
Tween	Non-ionique		1228 (-)	0.06 (0.0074)	Non	1%
Octyl glucoside	Non-ionique		292 (8K)	23-24 (~0.70)	Oui	1%
Digitonine	Non-ionique		1229 (70K)	<0.5		?
SDS	Anionique		288 (18K)	6-8 (0.17-0.23)	Non	0,05%
Sodium deoxycholate	Anionique		415	2-6	Oui	10%
CHAPS	Zwitterionique		615 (6K)	8-10 (0.5-0.6)	Oui	1%
CHAPSO	Zwitterionique		631 (7K)	8-10 (~0.505)	Oui	1%

**Tableau 2 : Exemple de détergents classiquement utilisés en biochimie et propriétés associées.**

D'autres facteurs peuvent influencer les résultats d'une expérience d'immunoprécipitation ou de pull-down tels que la durée de l'expérience [128], la température à laquelle est réalisée l'expérience ou bien la quantité de protéines dans le lysat. Par conséquent la réalisation d'une expérience de purification d'affinité nécessite d'avoir mûrement réfléchi au design de l'expérience qui doit également intégrer des contrôles appropriés. Au-delà de l'optimisation des conditions

expérimentales, différentes stratégies ont été élaborées pour essayer de discriminer les « vraies » interactions des interactions non-spécifiques.

Certaines stratégies consistent à éliminer les protéines abondantes généralement présentes dans les expériences contrôles [129] ou appartenant à des catégories de protéines considérées comme des contaminants classiques (protéines du cytosquelette, protéine de chocs thermiques, protéines ribosomales) [130]. La reproductibilité des résultats d'une expérience est souvent utilisée comme critère pour filtrer les listes de partenaires potentiels en considérant que les interactions non-spécifiques sont aléatoires et ne seront donc pas vues systématiquement dans les expériences. Cependant, ce type de stratégie peut entraîner l'élimination à tort de vrais partenaires d'interaction [131].

Il est ainsi apparu nécessaire de développer **des approches quantitatives pour filtrer de manière plus fine les résultats des expériences de purification d'affinité**. En 2006, Gavin *et al.* ont été parmi les premiers à se baser sur des critères chiffrés pour traiter les données de protéomique obtenues à l'issue d'expériences de purification d'affinité en tandem. Ils ont ainsi introduit la notion d'index de socio-affinité (*Socio-Affinity Index* SAI) pour quantifier la propension de protéines à s'associer. Cet index tient compte de la fréquence des protéines dans le jeu de données et a pour but de discriminer les vraies des fausses interactions qui impliqueraient des partenaires très abondants [132].

Dans une étude globale des complexes protéiques chez *saccharomyces cerevisiae* par purification d'affinité en tandem couplée à la spectrométrie de masse, Krogan *et al.* ont utilisé la notion de probabilité d'interaction, valeur basée sur la reproductibilité expérimentale de l'interaction ainsi que sur les scores obtenus pour l'identification des protéines basés sur leurs spectres de masse [133].

Depuis, des approches quantitatives ayant recours au marquage chimique ou métabolique ont vu le jour, ainsi que des approches quantitatives sans marquage dites « label-free » s'appuyant sur du « spectral counting » (nombre de spectres MS/MS d'un peptide) ou sur l'intensité des ions précurseurs (données MS). Pardo M. et Choudhary J.S. signent une revue de qualité avec une comparaison critique des différentes approches utilisées [134].

Parmi les outils disponibles pour filtrer et évaluer la probabilité d'avoir de vraies interactions, j'ai choisi de citer le **logiciel SAINT** (*Significance Analysis of Interactome*) [135] que j'ai eu l'occasion d'utiliser dans le cadre de ce travail de thèse (*cf.* chapitre 6). Ce logiciel a été mis au point au départ pour étudier l'interactome d'une kinase et traiter les résultats d'une expérience de purification d'affinité avec une seule étiquette [136]. Il a ensuite été généralisé pour des jeux de données de différentes tailles comportant ou non des contrôles négatifs et est également utilisable lorsqu'il n'y a qu'une seule proie.

SAINT permet d'assigner une valeur de probabilité de « vraie » interaction pour chaque protéine qui interagit avec la protéine d'intérêt à partir du nombre de spectres MS/MS acquis ayant permis d'identifier cette protéine. Cette valeur de probabilité, qui est normalisée par rapport à la longueur de la protéine, peut aussi être moyennée en utilisant plusieurs réplicats d'une expérience et en tenant compte des valeurs obtenues pour les contrôles négatifs.

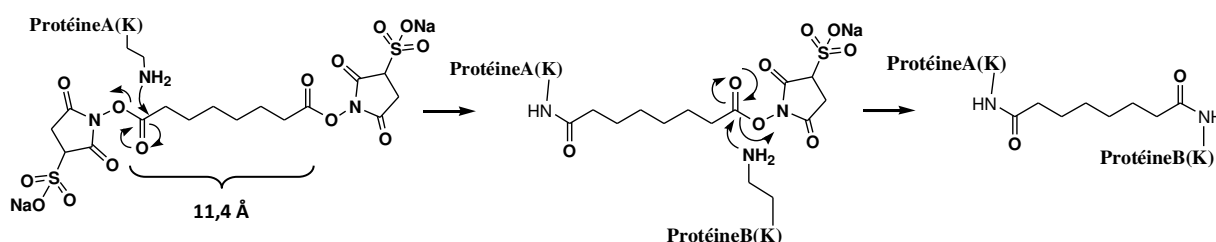
## 5. Pontage covalent ou "Cross-linking" couplé à la spectrométrie de masse

Les stratégies basées sur l'utilisation du cross-linking (pontage covalent en français) couplé à la spectrométrie de masse sont devenues de plus en plus courantes pour les études d'interactions protéine-protéine. Par commodité nous conserverons le terme anglais de cross-linking et ses dérivés pour la suite du manuscrit.

Le terme de cross-linking désigne le processus qui permet de relier par liaison chimique covalente deux molécules différentes. Le cross-linking peut être réalisé grâce à des réactifs de cross-linking de réactivité modérée ou *via* des intermédiaires hautement réactifs dans le cas du photocross-linking. Jusqu'à présent le cross-linking chimique est plus utilisé que le photocross-linking.

### a. Cross-linking chimique

Les réactions de cross-linking chimique sont généralement basées sur l'utilisation d'un cross-linker bifonctionnel (*cf.* exemple Figure 14) qui est constitué d'un bras espaceur composé d'une chaîne carbonée portant à chacune de ses extrémités des fonctions réactives qui peuvent être identiques (cross-linker homobifonctionnel) ou différentes (cross-linker hétérobifonctionnel). Les fonctions réactives sont principalement des esters activés ciblant soit les résidus lysine (par exemple les esters N-hydroxysuccinimide (NHS)) soit les résidus cystéine (ester maléimide). Cependant des réactions secondaires avec les résidus tyrosine, thréonine et sérine ont également été décrites [137].



**Figure 14 : Réaction de cross-linking chimique en deux étapes avec le cross-linker homobifonctionnel BS<sup>3</sup> qui réagit préférentiellement avec les amines des chaînes latérales des lysines.**

Le pH est le paramètre le plus critique : des valeurs entre 6,5 et 8,5 constituent un bon compromis entre la réactivité des lysines nucléophiles ( $pK_a \text{ } \epsilon\text{NH}_2$  9,5) et des cystéines ( $pK_a \text{ SH}$  8,3) envers les esters activés du cross-linker et le taux d'hydrolyse de ces derniers. De plus, une telle gamme de pH va permettre d'éviter la dénaturation de la protéine. Le tampon doit permettre de solubiliser le cross-linker et être dépourvu de groupements réactifs susceptibles d'interférer avec la réaction de cross-linking telles que les amines libres. La température, la concentration des partenaires d'interactions et du cross-linker ainsi que la longueur du bras espaceur (typiquement entre 4 et 30 Å) sont des paramètres qui peuvent également être optimisés.

Cependant le cross-linking chimique est souvent caractérisé par de faibles rendements (tout au plus quelques %) quelles que soient les fonctions réactives utilisées. De plus, une grande variété de produits de cross-linking est généralement obtenue conduisant à une dilution du signal. Afin de surmonter ces difficultés, des cross-linkers porteurs d'étiquettes d'affinité [138], [139] permettant un

enrichissement sélectif de l'échantillon en espèces cross-linkées ont été développés. L'utilisation de cross-linkers clivables [140], [141], marqués isotopiquement [142] ou substitués par une molécule de matrice CHCA pour l'analyse MALDI [143] améliore la détection ainsi que l'identification/caractérisation des espèces cross-linkées. La purification d'affinité *via* une étiquette portée par le cross-linker ou par un des partenaires d'interaction [144] et la technique SDS-PAGE sont les méthodes les plus couramment utilisées. Néanmoins, les méthodes de chromatographie échangeuses de cations [145], [146] ou d'exclusion stérique [147] constituent des approches prometteuses pour l'enrichissement d'espèces cross-linkées. L'analyse des peptides cross-linkés peu abondants nécessite l'utilisation de spectromètres de masse possédant une grande sensibilité (femtomolaire) pour leur détection et une grande précision en masse (inférieure à 10 ppm) pour leur identification dans la mesure où le nombre de combinaisons de deux peptides est énorme. De plus, la spectrométrie de masse en tandem est nécessaire pour la caractérisation des zones d'interaction. De nombreux logiciels dédiés à l'analyse d'espèces cross-linkées ont été développés de manière à réaliser une acquisition sélective et une interprétation spécifique des spectres de fragmentation. Parmi eux, FINDX a été conçu pour fragmenter sélectivement les cross-links entre protéines marquées  $^{14}\text{N}/^{15}\text{N}$  par LC-MALDI-TOF/TOF [148], xQuest [145] est dédié à la recherche de peptides cross-linkés marqués isotopiquement et Crosswork [149] ou Xlink-Identifier [150] sont destinés à l'interprétation d'analyses d'échantillons de cross-linking chimique sans marquage.

Les interactions transitoires telles que les interactions enzyme-substrat sont connues pour contribuer de façon importante à la dynamique de régulation des protéines. Cependant ces interactions transitoires sont difficiles à capturer par des expériences de purification d'affinité classiques. Les méthodes de cross-linking constituent des outils efficaces pour l'étude de ces interactions provisoires et ont fait leur preuve *in vitro*. Par exemple, Lambert *et al.* ont utilisé le cross-linker BS<sup>3</sup> pour mettre en évidence *in vitro* l'interaction ainsi que les résidus impliqués dans l'interaction de la protéine de choc thermique 21 (*Heat shock Protein Hsp21*) avec son substrat, la malate déshydrogénase (MDH) [151]. Le cross-linking chimique a également été intégré dans des expériences de purification d'affinité couplées à la spectrométrie de masse en tandem pour l'identification d'interactions transitoires ou faibles [152].

Cependant la plupart des cross-linkers chimiques ne sont pas utilisables pour des expériences *in vivo* car ils ne peuvent pas traverser efficacement la membrane plasmique des cellules mais surtout du fait de l'hydrolyse rapide de leurs groupements réactifs. Le paraformaldéhyde (PFA) et le glutaraldéhyde (cf. Figure 15) font figures d'exception puisqu'ils peuvent traverser la membrane plasmique et permettre de cross-linker des partenaires d'interaction *in cellulo*. Ils sont d'ailleurs beaucoup utilisés pour la fixation de tissus.

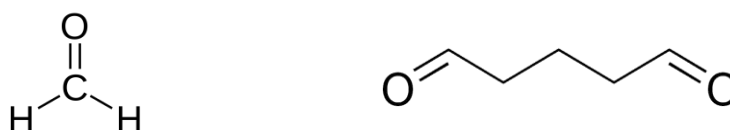
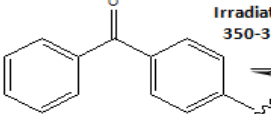
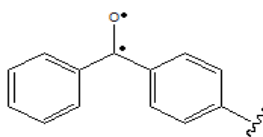
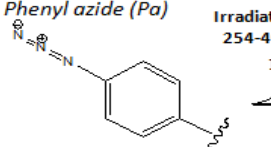
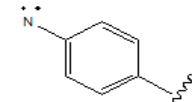
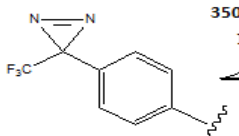
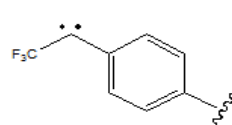


Figure 15 : Structure du formaldéhyde PFA (à gauche) et du glutaraldéhyde GA (à droite).

Le PFA, très petite molécule (4 atomes) correspondant à un bras espaceur de 2,3 à 2,7 Å, permet avant tout de maintenir des interactions transitoires qui seraient perdues dans des expériences de purification d'affinité classiques. Cependant, il ne permet pas d'étudier les zones d'interaction ni de savoir si l'interaction d'un partenaire identifié avec la protéine d'intérêt est une interaction directe ou non. En effet, la réaction de cross-linking avec le PFA étant réversible, les liaisons covalentes obtenues sont généralement détruites avant la digestion et l'analyse des partenaires d'interaction purifiés par immunoaffinité [153]. D'autre part, aucune analyse de spectrométrie de masse n'est parvenue à mettre en évidence l'existence de cross-links entre deux protéines ou deux peptides indiquant que la liaison covalente formée ne résiste pas au transfert dans le spectromètre de masse [154].

### b. Photocross-linking

Pour la capture d'interactions protéine-protéine *in vivo*, les efforts se sont donc concentrés sur la technique de photocross-linking. Le photocross-linking repose sur l'utilisation de groupements photoactivables pouvant, suite à une irradiation photonique, conduire à la formation d'une liaison covalente. Le tableau ci-dessous présente quelques exemples de groupements fréquemment rencontrés (Tableau 3).

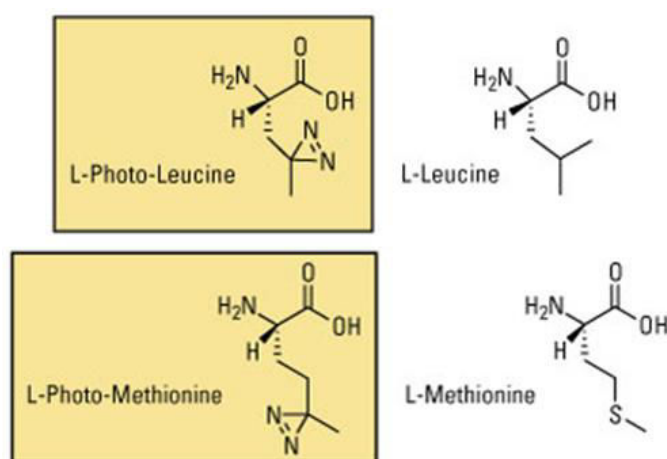
Précurseurs stables	Espèces hautement réactives	Réactivité
<b>Benzophenone (Bp)</b> 	Irradiation UV 350-365 nm 	Insertion du biradical au niveau des liaisons C-H
<b>Phenyl azide (Pa)</b> 	Irradiation UV 254-400 nm 	Insertion du nitrène au niveau des liaisons C-H et des hétéroatomes
<b>Trifluoromethyl phenyl diazirine (TFMFD)</b> 	Irradiation UV 350 nm 	Insertion du carbène au niveau des liaisons C-H et des hétéroatomes

**Tableau 3 : Trois groupements photoactivables communément utilisés : la benzophénone (Bp), le phényle azide (Pa) et le trifluorométhyl phényl diazirine (TFMFD).**

De la même manière que pour le cross-linking chimique, il existe des cross-linkers bifonctionnels portant généralement un groupement photoactivable à une extrémité et un groupement chimique réactif de l'autre côté. Cependant ces cross-linkers bifonctionnels sont plutôt réservés à des applications *in vitro* et peu, voire pas utilisés *in vivo* en raison de l'extraordinaire diversité des produits susceptibles d'être formés.

Plus souvent ces **groupements photoactivables** sont **directement portés par les biomolécules**. Ils sont soit **introduits lors de la synthèse** du peptide ou de la protéine en laboratoire (peptide ou

protéine qui pourra entrer dans la cellule ou y être vectorisé(e)), soit directement **incorporés dans les biomolécules au moment de la synthèse de ces dernières dans la cellule**. Cette synthèse endogène de protéines porteuses de groupements photoactivables peut être réalisée de différentes manières. En 2002, Chin J. et Schultz P. ont mis au point un ARNt ainsi que l'aminoacyl ARNt synthétase correspondant à cet ARNt permettant d'incorporer dans la machinerie de synthèse de protéines d'*Escherichia coli* l'acide aminé para-benzoyl-L-phenylalanine [155]. Quelques années plus tard, trois acides aminés ressemblant à l'isoleucine, la leucine et la méthionine mais contenant des cycles diazirine photoactivables et pouvant être incorporés tout à fait normalement dans la synthèse de protéines ont été synthétisés [156] (Figure 16).



**Figure 16 : Structure chimique des acides aminés photoactivables leucine et méthionine et de leurs équivalents naturels.**

Ensuite, une fois que la protéine photoactivable est « en place » dans son environnement cellulaire, une photoirradiation est réalisée et le (ou les) partenaire(s) d'interaction se situant à proximité sont liés covalamment. Finalement les cellules irradiées peuvent être lysées, les protéines d'intérêt purifiées et le (ou les) complexe(s) formé(s) identifié(s).

Les groupements photoactivables sont caractérisés par une forte réactivité du fait de l'obtention sous l'action d'un rayonnement UV, d'espèces intermédiaires radicalaires. Le caractère hautement réactif de ces radicaux conduit aux résultats suivants :

- 1) N'importe quelle molécule se trouvant à proximité, quelle que soit sa composition en groupements fonctionnels, pourra être liée de façon covalente. La réaction permet de créer une liaison C-C ou C-hétéroatome en fonction de la sonde utilisée.
- 2) Cette forte réactivité implique une demi-vie courte de l'intermédiaire réactif radicalaire (ns-μs) minimisant la probabilité de capture d'interactions non-spécifiques.
- 3) En l'absence de partenaires d'interaction à proximité, l'intermédiaire réactionnel reviendra à un état stable en réagissant avec de l'eau ou avec lui-même selon les cas et cela limitera également la possibilité d'identifier des interactions non-spécifiques.

Malgré les différents avantages décrits précédemment, le photocross-linking reste aujourd'hui encore assez peu utilisé pour plusieurs raisons. (1) L'incorporation des sondes photoactivables n'est pas une opération triviale et nécessite généralement des optimisations, par exemple, au niveau de la position de la sonde ou sur le pourcentage d'acides aminés photoactivables à incorporer ; (2) il est important de vérifier que la (ou les) sonde(s) photoactivable(s) ne vont pas venir modifier la structure tridimensionnelle de la protéine porteuse de la sonde et par conséquent perturber l'interaction de cette protéine avec ses partenaires ; (3) une des difficultés du photocross-linking réside dans le traitement des échantillons obtenus. Il faut en effet être capable de repérer (western blot) et d'isoler les espèces photocross-linkées (purification d'affinité) ; (4) **l'obstacle qui nous semble aujourd'hui le plus important dans le développement du photocross-linking est le manque de logiciels permettant l'interprétation des données** à la fois pour l'identification des partenaires mais aussi pour la détermination des sites de photocross-linking. C'est cela qui nous a incité à nous lancer dans le développement d'un logiciel capable d'interpréter les spectres de fragmentations d'espèces photocross-linkées en collaboration avec l'équipe de bioinformaticiens du Dr. Du X. (Université de Charlotte, Caroline du Nord, USA).

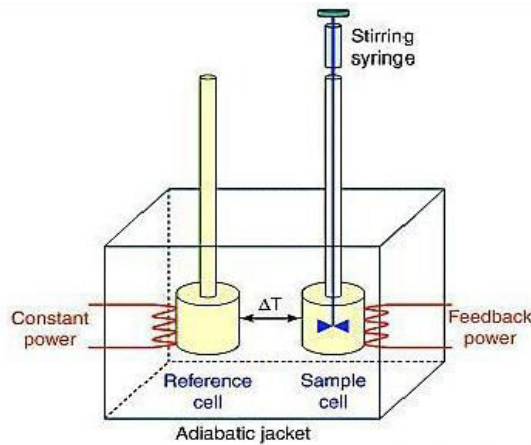
## II. Méthodes de validation et de caractérisation

Les méthodes de validation et de caractérisation qui sont présentées dans cette partie sont destinées à des expériences *in vitro*.

### 1. Calorimétrie de titrage isotherme (ITC)

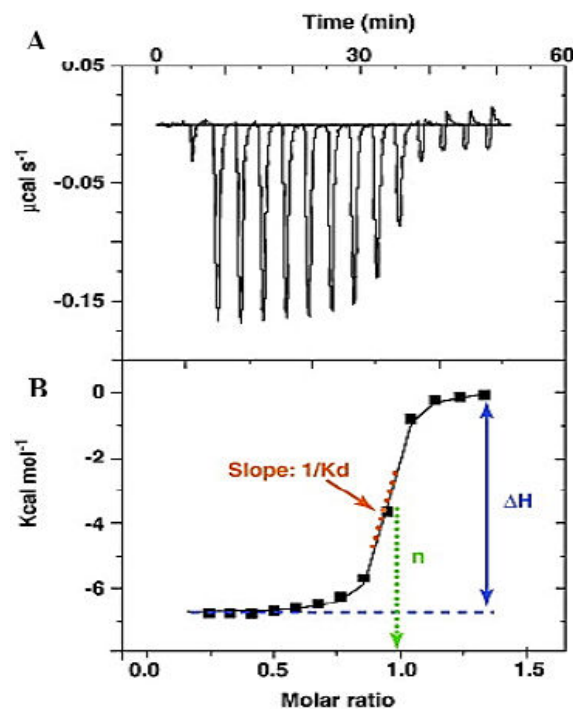
La calorimétrie par titrage isotherme (*Isothermal Titration Calorimetry* ITC) est une technique qui utilise les échanges de chaleur associés à une interaction moléculaire à température constante pour obtenir une caractérisation thermodynamique complète de cette interaction. En 1965, Christensen *et al.* ont été les premiers à décrire l'ITC comme une méthode permettant de déterminer simultanément la constante d'équilibre et l'enthalpie associée à la formation d'un complexe [157]. Ce système a ensuite été utilisé pour l'étude d'interactions au sein de systèmes biologiques [158]. Ce n'est qu'en 1989 que les premiers calorimètres commerciaux sont apparus.

Un système ITC est classiquement composé, comme schématisé Figure 17, de deux cellules, une cellule échantillon et une cellule de référence contenues dans une enveloppe adiabatique, ainsi que d'une seringue contenant un ligand qui est titré dans la cellule échantillon contenant la solution de protéine.



**Figure 17 :** Configuration classique d'un calorimètre de titrage isotherme (d'après Tanvi.D Site pharmaXchange.info).

Lorsque les deux éléments interagissent de la chaleur est dégagée (réaction exothermique) ou absorbée (réaction endothermique) dans la cellule échantillon. La puissance nécessaire pour maintenir les deux cellules à la même température est mesurée par le calorimètre et traduite en signal. Lorsque la protéine est saturée par le ligand ajouté, le signal de chaleur diminue jusqu'à ce que seule la variation de chaleur liée à la dilution soit observée. La représentation de ces variations de chaleur pour des ajouts de ligands réalisés à intervalles réguliers (Figure 18. A) permet de tracer une courbe représentant l'échange de chaleur en fonction du ratio molaire ligand/protéine. A partir de cette courbe la constante d'association ( $k_a = 1/k_d$ ), la stœchiométrie de la réaction de formation du complexe ( $n$ ) et la variation d'enthalpie associée peuvent être précisément déterminées (Figure 18.B). Avec ces mesures, il est ensuite possible de calculer l'entropie ( $\Delta S$ ) ainsi que changement d'énergie libre ( $\Delta G$ ) avec l'équation  $\Delta G = -RT \ln K_a = \Delta H - T \Delta S$ .



**Figure 18 :** (A) Données brutes extraites d'une expérience d'ITC. (B) Isotherme de liaison représentant l'échange de chaleur en fonction du ratio molaire ligand/protéine. (d'après Tanvi.D Site pharmaXchange.info).



L'ITC permet donc d'obtenir **une caractérisation thermodynamique complète d'une interaction**. Cela a conduit à un développement particulièrement rapide de la technique au cours des dernières décennies [159].

Aujourd'hui, l'ITC est utilisée pour étudier tous les types d'interactions dans lesquelles des protéines peuvent être impliquées : les interactions protéine-protéine, les interactions protéine-ADN/ARN et les interactions protéine-petite molécule. Parmi les applications de l'ITC on peut citer également l'étude du repliement et du dépliement de protéines et les études de titrage par déplacement d'un ligand permettant d'avoir accès à la constante d'association d'un ligand de haute affinité [160].

Contrairement à d'autres méthodes, l'ITC ne requiert pas l'immobilisation ou la modification chimique des protéines dans la mesure où l'absorption ou la production de chaleur sont des propriétés intrinsèques de quasiment toutes les réactions biochimiques. L'ITC présente néanmoins certains inconvénients puisqu'il s'agit d'une technique caractérisée par une faible sensibilité et qui nécessite par conséquent une consommation importante d'échantillon. C'est, de plus, une technique qui présente un bas débit, ne pouvant être utilisée en routine dans la mesure où les conditions doivent être optimisées pour chaque système étudié.

## 2. Résonance magnétique nucléaire (RMN)

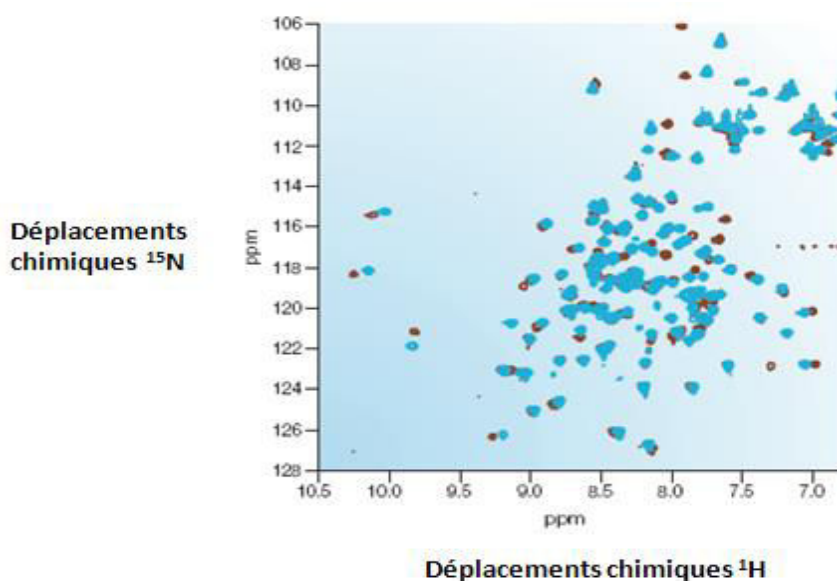
La RMN, technique basée sur l'interaction entre un champ magnétique et des noyaux atomiques à spin non nul, a été principalement connue pendant de nombreuses années comme méthode de détermination de structure tridimensionnelle de macromolécules biologiques. Aujourd'hui, c'est l'étude des interactions protéine-protéine qui donne lieu à la plupart des applications.

Jusqu'à présent la RMN n'a pas été en mesure de rivaliser avec la cristallographie aux rayons X en termes de résolution notamment. Cependant il se trouve qu'un certain nombre de complexes importants ne cristallisent pas correctement ou bien que la cristallisation ne se fait pas dans la conformation biologiquement active. Cela est particulièrement le cas pour les interactions faibles et/ou transitoires entre protéines. La RMN présente ainsi l'avantage par rapport à la cristallographie aux rayons X, de permettre l'étude des complexes en solutions. D'autre part, les avancées technologiques dans le domaine de la RMN ont permis de développer différentes méthodes d'étude des complexes qui fournissent des informations complémentaires pour la caractérisation d'interactions protéine-protéine.

Parmi elles, l'effet nucléaire Overhauser intermoléculaire (*Nuclear Overhauser Effect* NOE) qui permet de cartographier les interactions biomoléculaires. La technique NOE se base sur la mesure précise des distances entre les très nombreux protons d'un complexe pour établir la structure tridimensionnelle de ce dernier. Cette technique est seulement applicable lorsque l'interaction entre deux molécules est suffisamment forte ( $K_d < 10 \mu M$ ).

Nous pouvons citer également la technique de cartographie par perturbation du déplacement chimique (*Chemical shift Perturbation Mapping*) qui est la technique la plus utilisée pour définir les interfaces entre protéines. Brièvement, cette technique repose sur l'enregistrement de spectres HSQC (*Heteronuclear Single Quantum Coherence*)  $^{15}N$ - $^1H$  d'une protéine au fur et à mesure de l'ajout

d'un partenaire non marqué qui va venir modifier les déplacements chimiques de la protéine de départ. Un exemple de spectres HSQC pour ce type d'expérience est présenté Figure 19.



**Figure 19 : Exemple de spectres  $^1\text{H}$ - $^{15}\text{N}$  HSQC de l'interleukin-2 (IL-2) en présence (bleu) et en l'absence (marron) d'une petite molécule inhibitrice de IL-2 (Figure tirée de [161]).** La position de certains signaux est significativement modifiée suite à l'ajout de la petite molécule inhibitrice : Ces signaux correspondent à des acides aminés faisant partie de la zone d'interaction avec les inhibiteurs.

Cette méthode est sensible et permet de repérer des changements assez subtils. Par exemple, elle a permis de montrer que les domaines doigts de zinc des protéines peuvent, en plus de leur fonction de liaison de l'ADN, induire des interactions protéine-protéine avec de relativement faibles affinités ( $10^{-5}$  M) [162].

Un troisième exemple de technique faisant appel à la RMN pour la caractérisation de complexes est le titrage par RMN (*titration with NMR*). Cette technique, en plus de permettre de cartographier l'interface d'interaction, donne une bonne estimation de l'affinité, de la stœchiométrie, de la spécificité et des paramètres cinétiques de l'interaction. Un bon exemple d'application est le travail réalisé avec la troponine C qui est une protéine liant le calcium ainsi que la troponine I et qui est impliquée dans la régulation musculaire [163]. Le titrage par RMN a révélé que deux équivalents de calcium pouvaient lier la troponine C avec une haute affinité. Au cours de ce processus, la troponine C passe d'un état largement déstructuré à la formation d'un domaine capable d'interagir avec la troponine I. En présence de deux atomes de calcium, la troponine C lie un fragment de troponine I avec une stœchiométrie 1:1 et une constante d'affinité de  $2 \pm 1$   $\mu\text{M}$ .

Il a récemment été montré que la RMN pouvait permettre de mettre en évidence des changements conformationnels de l'ordre de la milliseconde [164].

Cependant, la RMN présente l'inconvénient de consommer des quantités importantes de matériel biologique, de l'ordre du milligramme par analyse, et ne permet d'analyser que des interactions avec une affinité faible de l'ordre du millimolaires ou de quelques micromolaires.

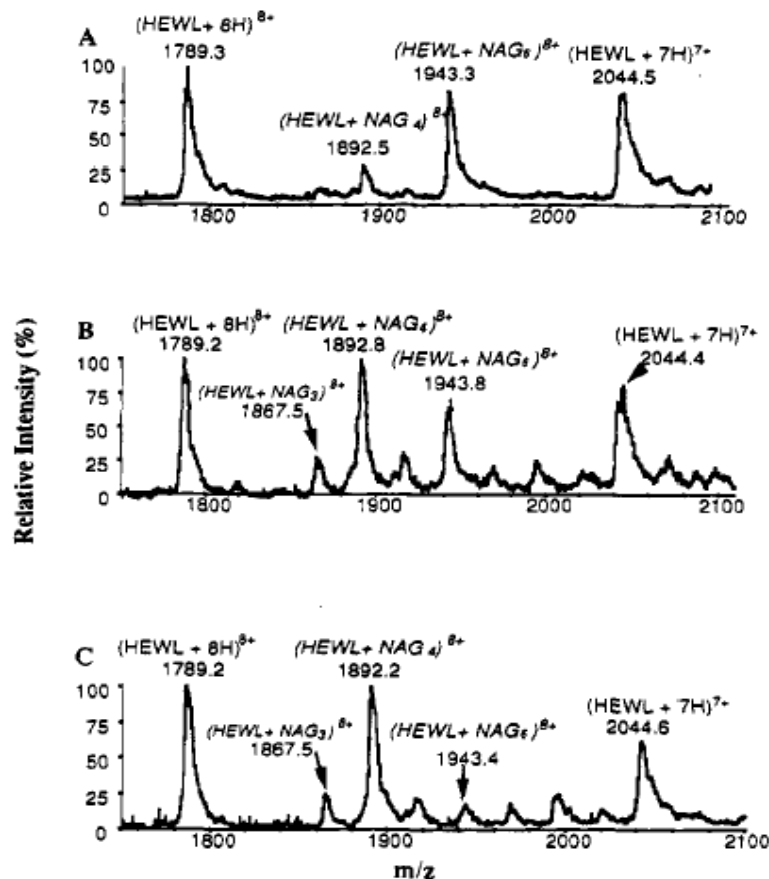
### 3. Cristallographie aux rayons X

La cristallographie aux rayons X est la technique la plus résolutive pour étudier la structure tridimensionnelle de biomolécules en phase solide. Cette technique permet, en effet, d'obtenir des informations structurales à l'échelle atomique. L'étude précise de la structure de ces molécules permet ainsi d'identifier les zones d'interaction entre différentes molécules mais aussi d'identifier d'éventuels sites de fixation de ligands de synthèse. Ainsi la cristallographie aux rayons X a été très utilisée dans l'industrie pharmaceutique pour la conception de ligands de haute affinité. Cependant elle s'est montrée à plusieurs reprises insuffisante pour correctement prévoir les relations structure-activité et doit ainsi souvent être couplée à des méthodes donnant accès à des informations thermodynamiques indispensables à une prédiction correcte des interactions [165]. L'obtention de cristaux diffractants et la nécessité de disposer de quantités importantes d'échantillons restent les principales limitations de cette technique [166].

### 4. Spectrométrie de masse native

Les techniques d'ionisation électrospray (ESI) et de désorption ionisation laser assistée par matrice (MALDI) sont des méthodes d'ionisation qualifiées de « douces » et permettent d'obtenir en phase gaz des ions de biomolécules de hautes masses moléculaires. Cependant, il apparaît que la technique MALDI est moins adaptée à l'étude d'interactions protéine-protéine et protéine-ligand dans la mesure où les conditions de matrice utilisées viennent perturber ces interactions. Différentes méthodes, telles que l'utilisation de conditions particulières pour l'obtention des cristaux de matrice, peuvent être mise en œuvre pour contourner cette difficulté et étudier des complexes non-covalents en MALDI [167]. Cependant aujourd'hui le terme de spectrométrie de masse native fait majoritairement référence à l'ESI.

Depuis 1991, la spectrométrie de masse avec une **ionisation de type électrospray** (ESI-MS) a été au cœur de travaux de recherche intenses visant à développer une application bien particulière, à savoir l'étude de complexes non-covalents. En spectrométrie de masse classique aussi appelée spectrométrie de masse moléculaire, les espèces ionisées sont présentes en phase gaz dans le spectromètre de masse sous forme individuelle après la destruction des liaisons non-covalentes existant en solution. En revanche la **spectrométrie de masse supramoléculaire** ou **native** a pour objectif de transférer les complexes moléculaires intacts, c'est-à-dire tels qu'ils préexistent en solution, vers la phase gaz de l'instrument. Au premier abord une technique nécessitant le transfert des espèces en phase gaz peut sembler inappropriée pour étudier des assemblages de molécules maintenus par des interactions faibles (liaisons électrostatiques, interactions de Van der Waals, liaisons hydrogènes et hydrophobes) et donc intrinsèquement fragiles. Cependant au début des années 1990 deux groupes américains ont été les premiers à montrer que des interactions protéine-ligand pouvaient survivre au processus d'ionisation électrospray [168], [169] (cf. Figure 20). A noter cependant que les interactions hydrophobes ne sont pas conservées lors de la désolvatation.



**Figure 20 : Spectres de masse obtenus en infusion avec une source électrospray et permettant de suivre l'hydrolyse de l'hexasaccharide de N-acétylglycosamine (NAG<sub>6</sub>) par le lysozyme de blanc d'œuf de poule (HEWL).** Le mélange de HEWL ( $3,17 \cdot 10^{-5}$  M) et NAG<sub>6</sub> ( $4,41 \cdot 10^{-4}$  M) dans le tampon acétate d'ammonium 10 mM à pH 5 est infusé dans la source électrospray à un débit de 2  $\mu$ L/min. Les spectres de masses présentés sont obtenus respectivement à 1 min (A), 10 min (B) et 60 min (C). L'abondance du pic correspondant au complexe (HEWL+NAG<sub>6</sub>)<sup>8+</sup> ( $m/z = 1943,3$ ) diminue en intensité au fur et à mesure du temps alors que celle du pic correspondant à (HEWL+NAG<sub>4</sub>)<sup>8+</sup> ( $m/z = 1892,2$ ), produit d'hydrolyse, augmente (Figure tirée de [168])

Depuis cette découverte, différents laboratoires ont travaillé à la mise au point de conditions expérimentales permettant de préserver la structure quaternaire des protéines et d'étudier la topologie des complexes et leur dynamique de manière reproductible.

La mise au point des conditions expérimentales commence par **le tampon**. Généralement, les tampons utilisés pour les purifications et les extractions de protéines ou de complexes non-covalents contiennent des sels non volatils (tampon phosphate, tris ou HEPES) qui ne sont pas compatibles avec l'ionisation électrospray. Ainsi la première étape, essentielle pour l'analyse de complexes non-covalents en spectrométrie de masse, est l'échange du tampon de purification lors d'une **étape de dessalage**. Le nouveau tampon doit remplir deux conditions : (1) être compatible avec l'ionisation électrospray et (2) permettre de préserver l'intégrité du complexe non covalent. Ce sont les tampons ammonium (acétate d'ammonium, carbonate d'ammonium, triéthylammonium) qui sont les plus adaptés. Ils permettent d'avoir un pH compris entre 5 et 8,5 et leur force ionique peut être adaptée en fonction de la stabilité des complexes en faisant varier leur concentration entre 5 et 500 mM.

Un autre point crucial pour maintenir les édifices non-covalents lors du processus de désorption et d'ionisation des complexes est **le réglage des paramètres du spectromètre de masse** qui contrôlent l'énergie communiquée aux ions lors du premier stade de pompage dans l'instrument. Cette région de pression intermédiaire est appelée interface et correspond physiquement à la zone des premiers hexapoles. Les paramètres à optimiser sont (1) la pression de l'interface  $P_i$  qui affecte l'efficacité des collisions et (2) le voltage d'accélération ( $V_c$ ) encore appelée tension cône qui contrôle l'énergie cinétique donnée aux ions dans la source de l'instrument [170]. Les paramètres  $V_c$  et  $P_i$  doivent être optimisés simultanément de manière à obtenir le meilleur compromis possible entre la désolvatation des ions, la transmission des ions de hauts  $m/z$  et la conservation des complexes non covalents.

Enfin, l'étude de complexes non-covalents par spectrométrie de masse native requiert l'utilisation d'un **analyseur capable de mesurer avec une bonne résolution de hautes valeurs de ratio  $m/z$** . Les instruments temps-de-vol (TOF : *Time-Of-Flight*) et les analyseurs hybrides quadripolaire-TOF (Q-TOF) sont particulièrement bien adaptés pour les expériences de spectrométrie de masse native car ils combinent haute sensibilité, haute résolution, vitesse d'acquisition et large gamme de masse (théoriquement illimitée) [171], [172].

Aujourd'hui la spectrométrie de masse native constitue un outil robuste qui est de plus en plus utilisé pour l'étude de complexes non-covalents de différentes natures (protéine/protéine, protéine/ligand, protéine/métal, protéine/ARN, protéine/ADN, etc...).

La spectrométrie de masse native ne permet pas d'obtenir d'informations détaillées sur la structure moléculaire et atomique des complexes, mais a un nombre important d'avantages par rapport aux techniques de biophysiques traditionnelles (RMN, calorimétrie de titration isotherme, cristallographie aux rayons X), telle que sa sensibilité (analyse de complexes protéiques disponibles en quantités picomolaires), sa rapidité, sa capacité à mesurer simultanément différentes espèces d'un mélange. Elle présente également l'avantage de pouvoir suivre en direct des changements réalisés en solution (ajout de ligand, ajout de protéine). Finalement elle permet d'obtenir des données capitales dans l'étude de complexes à savoir la stœchiométrie des complexes, la spécificité des interactions ainsi que les valeurs des constantes d'affinité relatives [173].

Ainsi la technique de spectrométrie de masse native est un outil de choix pour l'étude *in vitro* de complexes protéiques purifiés.

### 5. L'échange hydrogène/deutérium (HDX)

L'échange hydrogène/deutérium couplé à la spectrométrie de masse (*Hydrogen deuterium exchange-mass spectrometry* HDX-MS) est devenu une technique clé pour le suivi des changements dynamiques de structure des protéines en solution. La possibilité de réaliser des échanges hydrogène/deutérium pour étudier la structure des protéines a été mise en évidence par Kaj Ulrik Linderstrøm-Lang dans les années 1950 mais les premiers résultats d'étude de protéines par HDX-MS n'ont été publiés qu'en 1991 [174].

Cette approche repose sur le fait qu'une protéine en contact avec de l'eau lourde deutérée ( $D_2O$ ) va échanger ses atomes d'hydrogènes contre des deutériums. L'échange hydrogène/deutérium sera quasi-instantané pour les hydrogènes labiles, comme ceux des amines primaires par exemple. Les

protons amides du squelette protéique ont, eux, des cinétiques d'échanges plus lentes (de quelques minutes à quelques jours) qui varient en fonctions de la conformation et de la dynamique de la région de la protéine dans laquelle ils se trouvent. Ainsi, les hydrogènes des fonctions amides des régions désordonnées et en contact avec le solvant, c'est-à-dire ceux qui ne sont pas impliqués dans des liaisons hydrogènes stables, vont être échangés rapidement contre des deutériums. Par contre, ceux de zones bien repliées, associées entre elles de manière stable, vont être bien mieux protégés de ces échanges H/D.

La plupart des expériences HDX-MS sont réalisées en conditions de marquage continu (*continuous labelling strategy*) pour lesquelles le niveau de deutération des protéines est fonction du temps d'exposition au D<sub>2</sub>O. L'objectif de ces études est de détecter les modifications de structure induites par un stimulus externe telle qu'une modification chimique ou la liaison d'un ligand. Au bout de différents temps t (généralement de quelques minutes à quelques heures), l'échange H/D est stoppé par une acidification brutale du milieu (pH~2) et une baisse de la température (~0°C) afin de réduire l'échange inverse, c'est-à-dire le remplacement d'un deutérium d'une fonction amide par un hydrogène. Ensuite, une protéase fonctionnelle en conditions acides telle que la pepsine est ajoutée afin de réaliser une protéolyse limitée. Les peptides obtenus sont dessalés et séparés en chromatographie en phase inverse et analysés en spectrométrie de masse. Même en conditions acides et à basse température il est difficile d'éviter totalement l'échange inverse. Ainsi, il faut faire en sorte que la digestion et la séparation chromatographique soient réalisées en un minimum de temps. La réalisation d'analyses de spectrométrie de masse en tandem est nécessaire pour l'identification des peptides issus de la digestion à la pepsine qui réalise des coupures non spécifiques.

Ensuite le taux de deutération pour les différents peptides est calculé aux différents temps t de l'expérience selon la formule suivante:

$$\text{Taux de deutération}(t) = \frac{m(t) - m_0}{m_{100} - m_0}$$

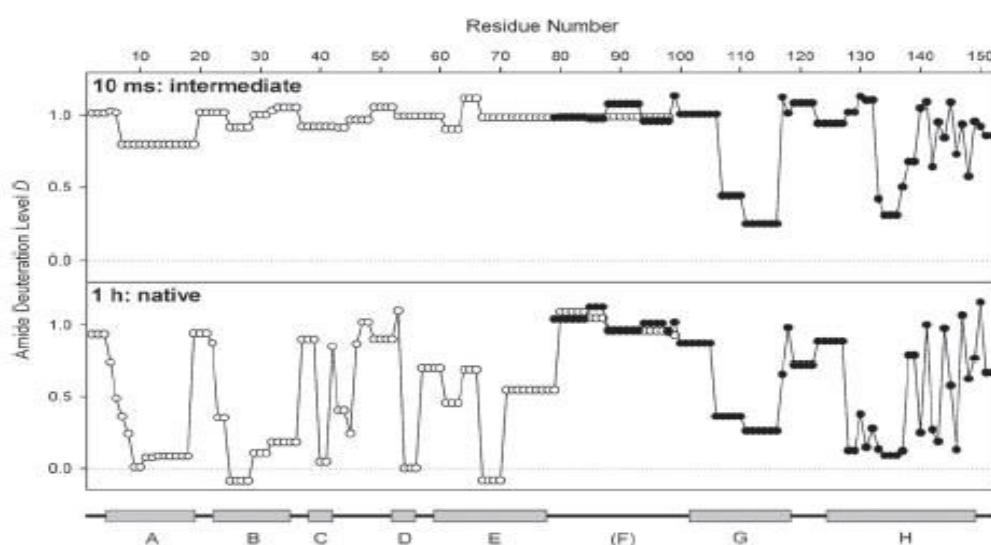
Avec m(t) la masse moyenne du peptide d'intérêt, m<sub>100</sub> et m<sub>0</sub> les masses du peptide pour des échantillons de référence respectivement complètement ou pas du tout marqués.

Différentes alternatives à ce protocole initial d'expérience HDX-MS basée sur une approche bottom-up ont rapidement été développées. Des expériences HDX-MS dites pulsées ont également été mises au point par Konermann *et al.* et sont utilisées pour détecter et caractériser des intermédiaires ayant une durée de vie courte (de l'ordre de la ms).

L'approche top-down est extrêmement prometteuse dans le domaine de l'HDX-MS [175] car elle permet de pallier les problèmes d'échanges inverses D → H pouvant avoir lieu pendant la protéolyse qui conduisent notamment à la perte d'information sur les groupements amines N-terminaux qui subissent systématiquement cet échange. Concernant la méthode de fragmentation, il faut savoir que la technique CID peut présenter des biais en induisant des processus de « scrambling » qui viennent complètement modifier l'empreinte H/D générée. Il faut par conséquent utiliser préférentiellement des techniques d'ionisation basées sur les électrons telles que l'ECD et l'ETD qui permettent de conserver l'information. Ce type d'expérience associant l'échange hydrogène-deutérium pulsé à la

spectrométrie de masse top-down avec fragmentation ETD a été mis en œuvre avec succès pour l'étude d'une forme intermédiaire de l'apomyoglobine existant lors de son repliement [176]. Les calculs des taux de deutération pour les différentes parties de la protéine ont permis d'obtenir le graphique ci-dessous pour des pulses à différents temps et ainsi suivre le repliement de la protéine (Figure 21).

Avec l'arrivée de systèmes robotisés permettant de réaliser en ligne toutes les étapes de l'expérience HDX-MS (en bottom-up ou en top-down), nous pouvons prévoir une augmentation future du nombre d'utilisateurs de la technique.



**Figure 21 :** Graphiques donnant le taux de deutération de différents domaines de la protéine apomyoglobine pour une expérience d'HDX-MS pulsée couplée à la spectrométrie de masse top-down avec fragmentation ECD pour des pulses réalisés à  $t = 10$  ms (en haut) et  $t = 1$  h (en bas) après l'induction du repliement (figure tirée de [176]).

## 6. Le marquage par oxydation (OFP)

Le marquage par oxydation couplé à la spectrométrie de masse (*oxidative footprinting mass spectrometry* OFP-MS) se rapproche beaucoup dans son principe de la technique HDX-MS. Il vise également à l'étude dynamique de la structure tridimensionnelle de protéines impliquées ou non dans des interactions protéine-protéine, protéine-ligand et repose également sur le marquage de groupements différemment accessibles en fonctions des conditions. Dans ce cas le marquage se fait au moyen de radicaux hydroxyles  $\text{OH}^\cdot$ . Ces radicaux peuvent être produits de différentes manières : par voie chimique avec la méthode de Fenton principalement [177], par photolyse de peroxyde ou par radiolyse de l'eau. Une fois générés, ces radicaux vont rapidement réagir avec les chaînes latérales des acides aminés situés à la surface des macromolécules. Lorsque la photolyse de peroxydes ou la radiolyse de l'eau sont réalisées avec un rayonnement laser ou synchrotron elles ont un avantage majeur sur la méthode chimique car elles permettent de produire des radicaux extrêmement rapidement et ainsi d'étudier la dynamique des protéines. En effet, en quelques



secondes un nombre suffisant d'oxydations va pouvoir être réalisé, oxydations qui seront ensuite détectées et quantifiées en spectrométrie de masse [178].

### **III. Approches de protéomique différentielle quantitative pour l'étude des réseaux d'interactions**

Aujourd'hui la spectrométrie de masse est une technique devenue incontournable dans le domaine de l'étude des interactions protéine-protéine. Elle occupe une place de choix à toutes les étapes de ces études, c'est à dire de la découverte de nouveaux partenaires d'interactions à la caractérisation complète de ces interactions (zones d'interaction, stœchiométrie, constante d'association etc...) comme nous avons pu le voir précédemment.

Au delà de l'identification des partenaires d'interaction d'une protéine d'intérêt, les études d'interactions protéine-protéine par spectrométrie de masse passent de plus en plus par des expériences de protéomique différentielle globale.

Les expériences de protéomique quantitative différentielle consistent à comparer les niveaux d'expression des protéines ou la nature et les niveaux des modifications post-traductionnelles de ces protéines dans des populations de cellules différentes ou ayant subi des traitements distincts. La stratégie la plus fréquemment utilisée est qualifiée de « shotgun » (cf. chapitre 2 partie I.3.b) et consiste à identifier et à quantifier le plus grand nombre de protéines possible par une approche bottom-up de manière à avoir une couverture importante du protéome.

Utilisées au départ majoritairement pour mettre en évidence des biomarqueurs de pathologie, ces expériences de protéomique quantitatives différentielles peuvent, grâce à l'intégration des données d'interaction dans des bases, permettre de comprendre comment des interactions protéine-protéine transmettent un signal et quels modules fonctionnels pourraient ainsi être visés pour contrer les dysfonctionnements observés dans de nombreuses pathologies.

Nous allons voir dans cette partie quelles sont les méthodes les plus utilisées en protéomique quantitative différentielle et dans quelle mesure elles peuvent compléter les études ciblées d'interactions protéine-protéine.

#### **1. Présentation des principales méthodes de quantification en protéomique différentielle**

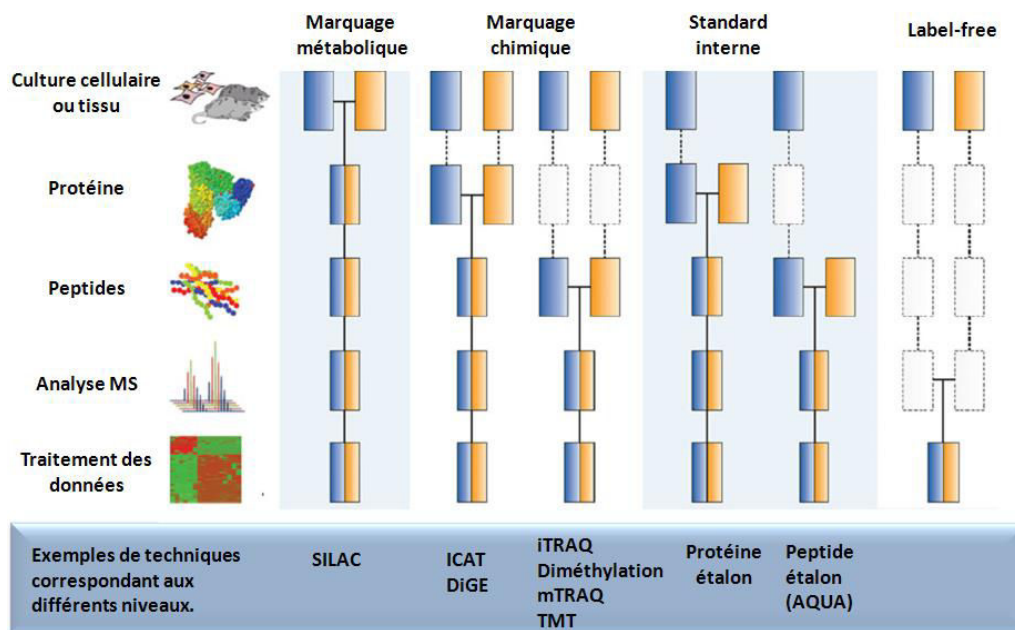
Les premières méthodes quantitatives de spectrométrie de masse mises au point reposent sur l'utilisation de marquage isotopique et ont été conçues au départ pour de la quantification absolue [179]. Les méthodes de quantification utilisées majoritairement aujourd'hui sont des méthodes de quantification relative et ont été développées au cours des vingt dernières années. Beaucoup s'appuient sur l'utilisation d'isotopes stables même si actuellement les stratégies de quantification sans marquage qualifiées de « label-free » connaissent de plus en plus de succès.

##### **a. Marquage métabolique :**

L'utilisation du marquage par des isotopes stables pour la quantification de peptides ou des protéines repose sur le fait que les propriétés physico-chimique des peptides naturels et marqués



isotopiquement, y compris leur comportement lors d'une analyse par spectrométrie de masse, sont très similaires voire identiques. Ainsi la quantification relative d'un échantillon par rapport à un autre peut être réalisée en comparant le signal MS des peptides marqués ou non. Dans le cas du marquage métabolique, les isotopes stables sont introduits dans chaque protéine au cours de la croissance cellulaire, ce qui permet d'avoir un standard marqué de chaque protéine d'un échantillon d'intérêt. Cela correspond également à l'étape de marquage la plus en amont dans un protocole d'expérience de protéomique, permettant ainsi d'avoir une grande précision et justesse dans la quantification en limitant les erreurs systématiques liées à la manipulation d'échantillon dans la suite du protocole (cf. Figure 22).

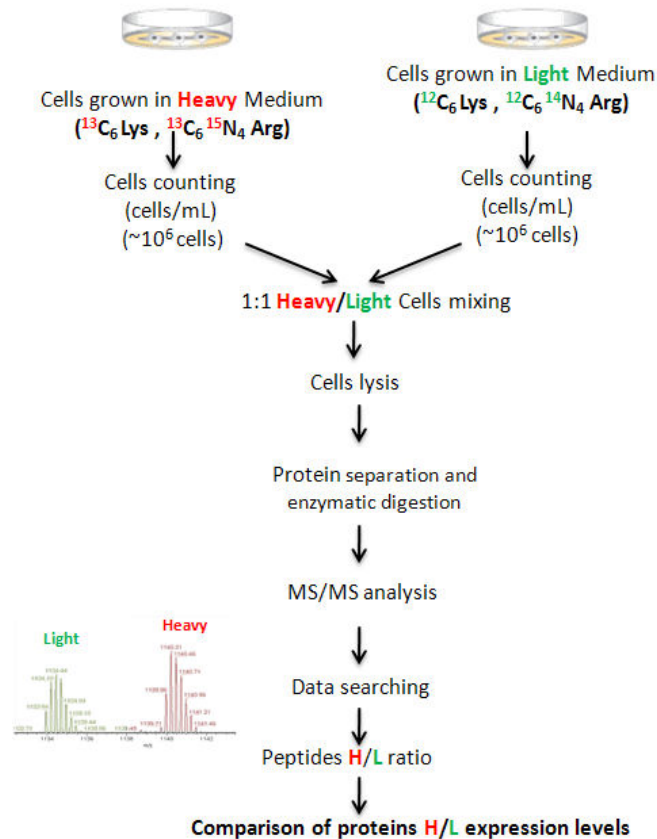


**Figure 22: Schéma des protocoles de spectrométrie de masse quantitative indiquant à quelle étape les deux conditions d'une expérience sont rassemblées et donc à quel niveau l'introduction d'erreurs systématiques dans la quantification liées à la manipulation d'échantillon est arrêtée.** Les rectangles bleus et jaunes correspondent aux deux conditions expérimentales à comparer. Les rectangles blancs délimités par des pointillés représentent les étapes au cours desquelles des variations expérimentales sont susceptibles d'introduire des erreurs dans la quantification. (Figure adaptée de [180])

Introduit il y a déjà une dizaine d'année par Ong *et al.* [95], la méthode de marquage métabolique la plus populaire est nommée méthode **SILAC** pour *Stable Isotope Labelling with Amino acids in Cell culture*. Dans les expériences de SILAC classique les acides aminés lysine et arginine naturels des milieux de culture sont remplacés par des acides aminés lysine et arginine marqués  $^{13}\text{C}$  et  $^{15}\text{N}$  de façon à ce que tous les peptides obtenus après digestion trypsique contiennent au moins un acide aminé marqué. Le schéma d'une procédure de SILAC classique est présenté Figure 23.

La quantification relative est réalisée en comparant les intensités des massifs isotopiques des peptides marqués et des peptides non marqués correspondant, visibles dans le spectre de masse.

## Introduction : Méthodes d'étude des interactions protéine-protéine



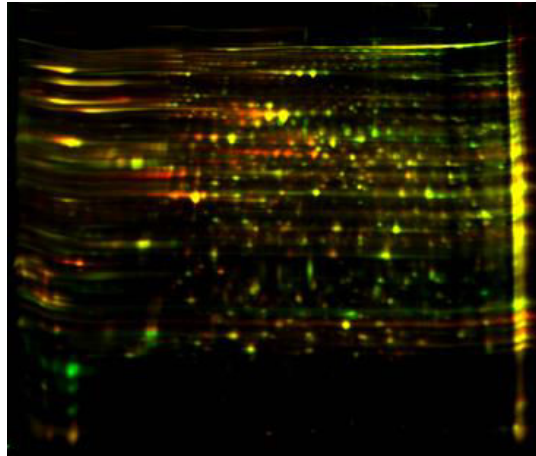
**Figure 23 : Exemple de protocole pouvant être utilisé pour réaliser la quantification relative de protéines par la technique SILAC.**

Il est possible de combiner plus de deux échantillons dans une même analyse. Cependant le répertoire d'acides aminés marqués restant encore limités, un maximum de trois échantillons sont classiquement comparés (par exemple Arg, Lys ;  $^{13}\text{C}_6$  Arg  $^{13}\text{C}_6$  Lys ;  $^{13}\text{C}_6$   $^{15}\text{N}_4$  Arg,  $^{13}\text{C}_6$   $^{15}\text{N}_2$  Lys). Relativement récemment une approche faisant appel aux cellules marquées SILAC a été développée pour la quantification de protéines dans les tissus. Dans cette approche appelée super-SILAC, un mélange de différentes lignées cellulaires marquées SILAC est utilisé comme standard interne pour la quantification des protéines dans les tissus [181], [182].

### b. Marquage chimique des protéines et des peptides

Une des plus anciennes techniques de protéomique quantitative, encore largement utilisée aujourd'hui, est l'électrophorèse sur gel différentielle ou *Differential Gel Electrophoresis* (DiGE). Cette technique introduite en 1997 par Unlü M. *et al.* [183] repose sur le marquage chimique d'échantillons de protéines correspondant à des conditions différentes par des fluorophores distincts. Les premiers fluorophores à avoir été développés sont des cyanines notées Cy2, Cy3 et Cy5 qui ciblent les résidus lysine. Ils ont des états de charge similaires, des poids moléculaires proches mais des propriétés spectrales absorption-émission suffisamment différentes pour les discriminer en utilisant des filtres optiques. Après marquage, les protéines des différentes conditions sont mélangées et subissent ensemble une migration sur gel d'électrophorèse 2D. Ensuite, en scannant le gel 2D obtenu aux longueurs d'ondes d'excitation des différents fluorophores, il est possible

d'identifier des variations d'expression des protéines. La Figure 24 est un exemple d'image de gel 2D DiGE résultat de la superposition des images obtenues pour deux échantillons marqués avec des fluorophores verts et rouges.



**Figure 24 : Image d'un gel 2D réalisée avec la technique DiGE.** L'échantillon contrôle est en vert et l'échantillon traité est en rouge. La couleur jaune correspond à la superposition de fluorophores rouges et verts. (Image tirée du site <http://www.appliedbiomics.com>)

Parmi les techniques de marquage chimique les plus populaires actuellement se trouvent la technique dite TMTs pour *tandem mass tags* et la technique iTRAQ pour *isobaric tags for absolute and relative quantification* qui toutes deux ciblent les résidus lysines et font appel à des étiquettes isobariques [184], [185]. Ces réactifs sont construits de manière à ce que les peptides porteurs d'étiquettes aient tous la même masse (isobarique) mais qu'ils puissent être distingués suite à leur fragmentation dans un spectromètre de masse grâce à un ion rapporteur (*reporter ion*) visible dans la région des basses masses (zone du spectre dans laquelle aucun autre signal ne vient interférer) des spectres MS/MS obtenus.

La comparaison de la quantité des peptides/protéines issus des différents échantillons est réalisée en comparant l'intensité relative des ions rapporteurs. Contrairement à d'autres techniques de marquage, l'utilisation d'étiquettes isobariques est bien adaptée aux analyses multiplexées de différents échantillons biologiques car (1) les réactifs sont disponibles dans 8 versions différentes, (2) la complexité de la séparation chromatographique n'est pas augmentée du fait de la co-élution des différents peptides marqués et (3) la complexité des spectres de masse n'est pas augmentée car les peptides sont isobariques.

En plus des méthodes de marquage chimiques par des étiquettes isobariques, deux autres méthodes de marquage chimique méritent d'être mentionnées ici.

La première est le marquage mTRAQ (*Mass differential Tags for Relative and Absolute Quantification*) qui utilise la même chimie que le marquage iTRAQ mais qui consiste à ajouter des étiquettes aux peptides de façon à ce qu'ils puissent être distingués en MS et pas seulement en MS/MS [186]. Un article de revue indique cependant que le marquage iTRAQ est meilleur que le marquage mTRAQ pour la quantification globale, à grande échelle des protéines et des phosphopeptides [187].

La seconde méthode est le marquage par diméthylation qui offre une alternative pour le marquage chimique de digests de protéines [188]. De la même façon que pour les méthodes précédentes, toutes les amines primaires (en N-terminal et sur la chaîne latérale des lysines) sont marquées.

Le marquage par diméthylation peut être réalisé avec du formaldéhyde contenant un ou deux atomes de deutérium et éventuellement également un  $^{13}\text{C}$  ( $\text{CH}_2\text{O}$ ,  $\text{CD}_2\text{O}$ ,  $^{13}\text{CD}_2\text{O}$ ) en présence de  $\text{NaBH}_4$  deutéré ou non. Il permet donc de comparer jusqu'à trois conditions en parallèle [189].

Le marquage par diméthylation présente de nombreux avantages :

- 1) Les réactifs nécessaires sont très peu coûteux.
- 2) Le marquage peut être réalisé aussi bien en solution que sur colonne de chromatographie en phase inverse (ou sur SPE).
- 3) Des quantités allant de quelques microgrammes à des milligrammes peuvent être marquées en suivant un même protocole.

Un inconvénient cependant est que les deutériums utilisés dans les étiquettes peuvent induire une légère modification du temps de rétention des composés lors de la LC-MS/MS dans la mesure où ils présentent une affinité légèrement différente pour la phase stationnaire. La réussite de ce type d'expériences reposant sur un marquage chimique implique, bien entendu, d'avoir des rendements très élevés.

### c. Quantification en « label-free »

La quantification par une méthode de « label-free » se fait, comme son nom le suggère, sans avoir recours à un marquage préalable des peptides ou des protéines.

Cette méthode de quantification des protéines qui connaît une expansion relativement rapide dans le domaine de la protéomique se fait essentiellement de deux manières : (1) en utilisant **directement l'intensité du signal** lié aux peptides tryptiques d'une protéine donnée ou (2) en déduisant la quantité de protéines de manière **indirecte** en se basant sur **le nombre de peptide-to-spectrum matches (PSMs)** aussi appelé *spectrum count* qui correspond au nombre de spectres MS/MS réalisés sur les peptides tryptiques de cette protéine et ayant conduit à l'identification de cette dernière.

Les approches de quantification en label-free basées sur l'intensité du signal MS des peptides passent par la production d'un chromatogramme d'ion extrait (XIC pour *eXtracted Ion Chromatogram*) qui va permettre de connaître la contribution d'un ion et donc d'un peptide donné au signal MS total. Bondarenko *et al.* ont montré qu'il y avait une excellente linéarité sur quatre ordres de grandeur entre l'intensité du signal dû à un peptide et sa concentration [190]. Des opérations d'alignement des signaux LC-MS/MS et la normalisation de l'intensité des pics sont cependant généralement nécessaires pour assurer une bonne quantification.

Les approches de comptage spectral sont basées sur l'observation que le nombre de PSMs, le nombre de peptides identifiés et la couverture de séquence obtenue pour une protéine donnée dans une expérience de LC-MS/MS sont corrélés avec la quantité de cette protéine. La quantification relative de l'abondance de protéines dans différents échantillons est réalisée directement en se

basant sur le nombre de PSMs obtenus pour chaque protéine identifiée, information qui est généralement donnée dans les tableaux de résultats des moteurs de recherches en banque de données. Cependant, il faut être conscient que ce type d'approche dépend des paramètres expérimentaux tels que les critères d'exclusion dynamique des ions précurseurs, la largeur des pics chromatographiques et la vitesse de scan du spectromètre de masse. Dans la mesure où en quantification label-free tous les échantillons à comparer sont analysés séparément, il faut veiller à ce que les conditions d'analyses LC-MS/MS soient stables et comparables pour tous ces échantillons (Reproductibilité des profils chromatographiques, stabilité de la source d'ionisation et de la détection). **Il faut donc un système extrêmement robuste pour garantir des résultats fiables.**

## 2. Intérêt pour les études d'interactions protéine-protéine

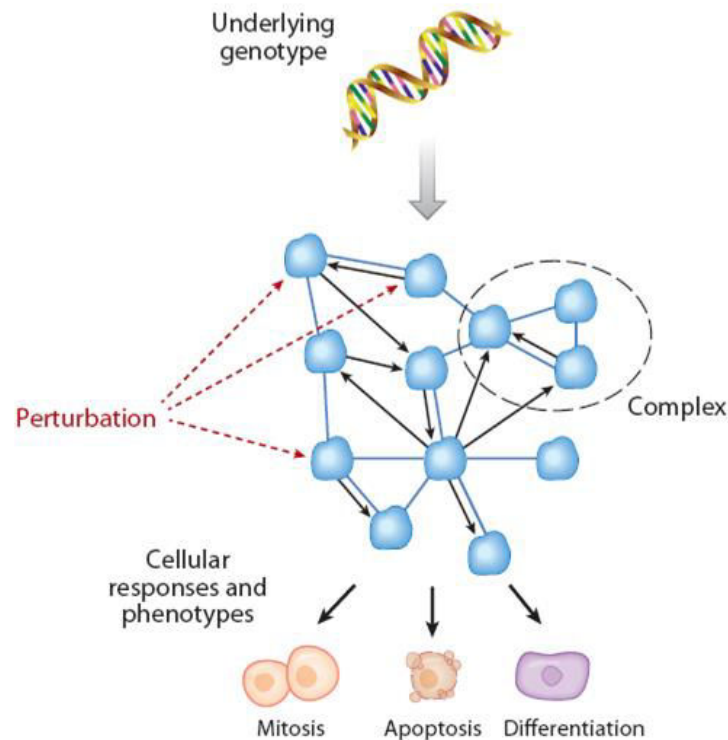
De nombreux projets ont permis de couvrir une part importante des protéomes et des PTMs chez différents organismes ainsi que de mesurer des changements dans l'abondance de ces protéines et de leurs PTMs [191], [192].

Dans la plupart de ces études de protéomique quantitative différentielle, le résultat principal est une liste de protéines dont le niveau d'expression a été modulé ou une liste de peptides qui ont été modifiés. Ensuite, typiquement quelques éléments de ces listes sont étudiés par des approches classiques de biochimie ou de biologie moléculaire. Cette stratégie peut donner des résultats intéressants mais n'est pas vraiment satisfaisante dans la mesure où la plupart des informations collectées ne sont pas utilisées. Une stratégie alternative existe néanmoins. Elle consiste à confronter les listes de protéines générées à des bases de données de voies de signalisation ou d'interactions protéine-protéine et à les soumettre à des **outils d'enrichissements de termes de *gene ontology***. D'après le site du Gene Ontology Consortium (<http://www.geneontology.org>), le terme de *Gene ontology* fait référence à la nomenclature mise au point dans le cadre du Gene Ontology Project. Ce projet a été développé par les responsables de trois bases de données (Flybase, Saccharomyces Genome database et Mouse Genome Database) de manière à donner des descriptions constantes aux produits de l'expression de mêmes gènes existants dans différentes bases de données. Ils ont pour cela élaboré trois classes de vocabulaires contrôlés (ontologies) qui décrivent les produits des gènes en fonction (1) des processus biologiques, (2) du compartiment cellulaires et (3) de la fonction moléculaire dans lesquels ils sont trouvés ou impliqués, et cela indépendamment de l'espèce.

DAVID *Database for Annotation, Visualization and Integrated Discovery* [193], [194] et STRING *Search Tool for the Retrieval of Interacting Genes/Proteins* [195], [196] sont deux exemples très utiles de logiciels gratuits avec une interface web qui permettent de soumettre des listes de protéines et de les confronter avec différentes bases de données. L'objectif de ces analyses est (1) de relier le contenu de la liste de protéines obtenues aux connaissances existantes sur les fonctions et les voies de signalisations dans lesquelles sont impliquées ces protéines ou de (2) **visualiser l'ensemble des protéines sous la forme d'un réseau d'interaction.**

Ces réseaux d'interactions de protéines se placent entre le génotype et le phénotype (*cf.* Figure 25). En effet, on peut considérer que la structure et la topologie de ces réseaux sont l'expression de l'information génétique, qu'ils sont modulés de façon dynamique à différentes échelles de temps par

des perturbations externes (ex : facteurs environnementaux) ou internes (ex : altérations génétiques) et que l'ensemble des propriétés du réseau d'interaction détermine le phénotype [197].



**Figure 25: Schéma du concept de biologie en réseaux d'interactions (Network biology).** Les réseaux d'interactions de protéines constituent un niveau intermédiaire entre le génotype et le phénotype et les perturbations appliquées à ce réseau sont à l'origine des variations du phénotype (Figure tirée de [197]).

Ainsi avec les approches de **protéomique différentielle quantitative**, on ne va pas directement sonder les interactions protéine-protéine mais plutôt essayer de **déduire des informations sur les relations existant entre les protéines sur la base de modifications observées suite à l'application d'une perturbation du système.**

#### IV. Complémentarité des différents outils d'étude des interactions protéine-protéine

Un tableau récapitulatif plutôt qu'un long discours : c'est cette option que j'ai choisie pour cette partie (cf. Tableau 4).

Il faut garder à l'esprit qu'il est fondamental de bien comprendre les limites des différentes techniques et la nécessité d'adapter les conditions expérimentales en fonctions de la problématique. Dans tous les cas il faudra, bien entendu, savoir préparer les échantillons de manière adéquate, prévoir des expériences contrôles adaptées et garder un esprit critique sur les résultats!



Tableau 4 : Avantages et inconvénients des différentes techniques d'étude des interactions protéine-protéine.

Méthodes d'étude IPP	Découvertes nouveaux partenaires	Caractérisation thermodynamique de l'interaction + Stoechiométrie	Identification zones d'interaction	Principaux avantages	Principaux inconvénients/difficultés
Double hybride	++	-	-	- Adapté à l'analyse haut-débit - Sensibilité - Expérience menée <i>in cellulo</i>	- Interaction testée dans le nucléoplasme - Interaction binaire uniquement
Résonance plasmonique de surface	+	+++	-	- Sensibilité - Faible consommation en échantillon - Pas de marquage préalable des protéines	- L'identification des partenaires d'interactions passe par un couplage à la spectrométrie de masse.
AP-MS	+++	-	-	- Travail avec des milieux biologiques complexes - Mise en évidence d'interactions à plusieurs partenaires	- Interprétation des résultats
Cross-linking	+	-	+	- Faible consommation d'échantillons - Facilité de mise en œuvre	- faible rendement des réactions - manque d'outils pour l'analyse automatisée (en particulier photocross-linking)
ITC	-	+++	-	- Pas d'immobilisation ou de modification chimique des protéines	- Forte consommation d'échantillon - Bas débit
RMN	-	-	++	- Diversité des informations obtenues	- Forte consommation d'échantillon Gamme limitée de constante d'affinité
Cristallographie	-	-	+++	- Haute résolution des structures	- Obtention de cristaux diffractants - Consommation d'échantillon - Etude en phase solide uniquement
Spectrométrie de masse native	+	+++	-	- Faible consommation d'échantillons - Possibilité de faire évoluer le système en cours d'analyse (dynamique) - Informations sur la stoechiométrie	- Préparation d'échantillon qui nécessite beaucoup de précautions.
HDX-MS	-	-	+++	- Suivi dynamique	- Précaution à prendre pour limiter échanges D→H (pH ≈ 2.5 Tz 0 °C)

Dans la mesure où les expériences de protéomique différentielle ne permettent pas directement de sonder des interactions protéine-protéine, j'ai fait le choix de ne pas les intégrer au tableau.

Malgré cela, les expériences de protéomique différentielle sont tout à fait complémentaires des expériences ciblées et ce sont les allers-retours constants entre ces différentes approches qui permettent *in fine* une meilleure compréhension des mécanismes moléculaires à l'origine d'un changement de phénotype.

### V. Perspectives : Vers une étude dynamique possible des interactions protéine-protéine, exemple de la technique AP-SWATH

Les complexes moléculaires responsables de fonctions transitoires telles que la transmission d'un signal sont sujets à de fréquents remaniements. Jusqu'à présent la plupart des études réalisées sur la composition des complexes présentent une vue figée du système. Afin d'acquérir une meilleure compréhension des processus biologiques qui sont dynamiques par nature, il est essentiel de **développer des méthodes capables de quantifier les changements du système au cours du temps**. C'est dans cet objectif de quantifier les changements dynamiques des réseaux d'interactions de protéines qu'a été développé l'outil AP-SWATH (*Affinity Purification combined with Sequential Window Acquisition of all Theoretical spectra*).

Cette technique AP-SWATH a été introduite peu de temps après la mise au point de la stratégie SWATH qui repose sur la réalisation de cycles d'analyse systématique de tous les ions (*Data independent Acquisition* DIA) sur des gammes de masses successives larges de quelques dizaines de  $m/z$  [198]. L'avantage marquant est l'identification de protéines/peptides minoritaires qui dans les techniques plus classiques (DDA) ne seraient pas fragmentés et donc pas identifiés. La technique AP-SWATH a été utilisée par Collins B.C et *al.* pour étudier la dynamique de l'interactome de la protéine 14-3-3 $\beta$  après une activation avec l'IGF1 (*insulin growth factor 1*) de la voie de signalisation insuline-PI3K-AKT. Les niveaux d'expression de près de 2000 protéines dont une centaine de partenaires de la protéine 14-3-3 $\beta$  ont été suivis après 1, 10, 30 et 100 min de traitement et ont permis d'obtenir des **informations sur la réorganisation dynamique de l'interactome** de la protéine 14-3-3 $\beta$  suite à cette perturbation [199].

Dans le même temps, Lambert et *al.* ont suivi les modifications induites par le traitement avec un inhibiteur de l'HSP90 sur l'interactome de CDK4 (cyclin dependent kinase) [200]. Ils ont également proposé une procédure d'étude de changements des interactomes pouvant être généralisée à l'ensemble des études en biologie des systèmes.

Ainsi nous pouvons facilement imaginer qu'avec le développement rapide de l'instrumentation en spectrométrie de masse et des ressources bio-informatiques associées, ce type d'approche pourra devenir un « grand classique » parmi les méthodes d'étude d'interactions protéine-protéine.



## **Conclusion du chapitre**

Ainsi cette partie, qui peut être considérée comme un inventaire, non-exhaustif bien sûr, de la « boîte à outils » du chercheur allant à la « pêche » aux partenaires d'interactions protéiques, vise également à mettre en avant les avantages et les inconvénients des différentes techniques et à voir dans quelle mesure elles sont complémentaires.

C'est dans cette boîte à outils que nous sommes venus piocher pour étudier l'effet observé du peptide vecteur (R/W)<sub>9</sub> sur les cellules tumorales EF. Parmi les méthodes décrites, certaines avaient déjà été utilisées pour caractériser l'interaction du peptide (R/W)<sub>9</sub> avec l'actine : c'est le cas le l'ITC, de la RMN ou encore du cross-linking chimique. Nous avons fait le choix d'une approche sans *a priori* pour la recherche de partenaires du peptide et avons donc été amené à croiser les informations obtenues par différentes méthodes : la purification par affinité couplée à la spectrométrie de masse, le photocross-linking ou encore une approche de protéomique différentielle globale basée sur la technique de quantification SILAC (*Stable Isotope Labeling in Cell culture*). Par conséquent avant de rentrer dans le vif de notre sujet, il m'a semblé utile de refaire le point sur les méthodes choisies pour étudier les interactions protéine-protéine mais aussi sur celles que nous aurions pu choisir et/ou que nous utiliserons peut-être à l'avenir.



## Chapitre 3 : Les peptides vecteurs

### I. Présentation générale des peptides vecteurs :

#### 1. Découverte des CPPs

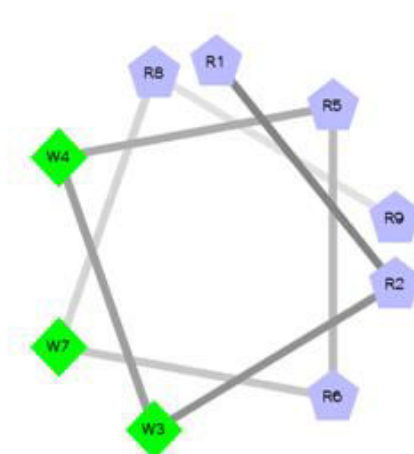
C'est à la fin des années 1980, qu'a été reportée pour la première fois la capacité d'internalisation d'une protéine dans les cellules contredisant l'idée établie que la membrane plasmique constituait une frontière infranchissable pour les molécules hydrophiles. En effet, en 1988, il a été démontré par deux équipes en parallèle que la protéine Trans-Activatrice de la Transcription (Tat) du virus HIV était capable de rentrer efficacement dans des cellules en culture et d'induire l'expression d'un gène viral [201], [202]. Plus tard, il a été montré que sur les 86 acides aminés que compte la protéine Tat, seul le domaine allant des résidus 48 à 60 (séquence CGRKKRRQRRRPPQC) et contenant une majorité d'acides aminés basiques était nécessaire à l'internalisation et à la translocation dans le noyau [203]. Le peptide correspondant à cette région de la protéine Tat et responsable de son internalisation a été nommé peptide Tat.

D'autre part, il a été découvert que la séquence de 60 aminés correspondant à l'homéodomaine Antennapédia de la drosophile était également capable d'entrer dans des cellules, dans ce cas, il s'agissait de neurones [204]. De la même façon que pour la protéine Tat, la capacité d'internalisation de l'homéodomaine a pu être attribuée à une région bien déterminée comprenant les résidus 43 à 58 et correspondant à la troisième hélice de l'homéodomaine [205]. Ce peptide de 16 résidus de séquence ERQIKVFQNRMRKWKK a été nommé Pénétratine.

Depuis d'autres peptides capables d'entrer spontanément dans les cellules ont été découverts ou mis au point de manière rationnelle : ils sont qualifiés de peptides vecteurs ou de CPPs pour *Cell Penetrating Peptides*.

#### 2. Les caractéristiques des CPPs et leur classification

De manière générale, les CPPs sont de courtes séquences peptidiques (moins de 30 acides aminés) capables d'entrer dans tous les types cellulaires avec une toxicité très limitée *via* des mécanismes dépendants ou indépendants de l'ATP et sans avoir recours à un récepteur spécifique. Beaucoup de CPPs sont nettement chargés positivement du fait de la présence dans leur séquence de plusieurs acides aminés basiques (arginine et/ou lysine). Cependant, la présence d'acides aminés acides n'est pas rédhibitoire tant que le nombre de charges négatives reste bien inférieur au nombre d'acides aminés basiques comme dans le cas du peptide VT5 [206]. Souvent les acides aminés basiques alternent avec des résidus hydrophobes comme le tryptophane. Cela confère aux CPPs (à l'exception des CPPs polycationiques purs comme R<sub>9</sub> composé de 9 résidus arginine) un caractère amphipathique et les conduits à se structurer en hélice amphiphile avec une face concentrant les résidus chargés et une face plus hydrophobe (cf. Figure 26).



**Figure 26 : Roue d'Edmunson : projection sous forme d'hélice alpha formée de la séquence (R/W)<sub>9</sub>. Les résidus hydrophobes sont en vert tandis que les résidus hydrophiles sont en bleu** (<http://rzlab.ucr.edu/scripts/wheel/wheel.cgi>).

Les CPPs sont généralement classés dans trois groupes en fonction de leur origine (cf. Tableau 5). Le premier groupe rassemble les **séquences peptidiques directement issues de protéines**. Les CPPs de ce premier groupe sont ainsi également appelés *Protein Transduction Domain* (PTD). Le deuxième groupe correspond aux **peptides chimériques** mis au point en associant des séquences pénétrantes naturelles. Enfin, le troisième groupe est celui des **peptides synthétiques** conçus de manière rationnelle suite aux observations réalisées sur les CPPs de séquences naturelles.

**Tableau 5 : Origines et séquences des CPPs les plus étudiés (adapté de [207])**

Classe	Peptide	Origine	Séquence	Référence
<b>Dérivés de protéines</b>	Pénétratine	Antennapedia (43-58)	RQIKIWFAQNRRMKWKK	[205]
	Peptide Tat	Tat (48-60)	GRKKRRQRRRPPQ	[203]
	pVEC	Cadherin (615-632)	LLIILRRRIRKQAHASK	[208]
<b>Chimériques</b>	Transportan	Galanine/Mastoporan	GWTLNSAGYLLGKINLKALAALAKKIL	[209]
	MPG	HIV-gp41/SV40T-antigen	GALFLGFLGAAGSTMGAWSQPKKKRKV	[210]
	Pep-1	HIV-reverse transcriptase/ SV-40 T-antigen	KETWWETWWTEWSQPKKKRKV	[211]
<b>Synthétiques</b>	Polyarginines	Inspirés de Tat	(R) <sub>n</sub> ; 6 < n < 12	[212], [213]
	MAP	<i>de novo</i>	KLALKALKALKAAALKLA	[214]
	(R/W) <sub>9</sub>	Inspiré de la Pénétratine	RRWRRRWRR	[215]

## **II. Mécanisme d'internalisation et outils d'études**

La question des mécanismes utilisés par les CPPs pour entrer dans les cellules, posée depuis leur découverte il y a maintenant plus de deux décennies, fait encore aujourd'hui l'objet de nombreux débats.

Les premières études réalisées avec la Pénétratine, le peptide Tat et R<sub>9</sub> ont conclu à une internalisation par **translocation directe** c'est à dire par une perturbation physique de la membrane plasmique indépendante de la température et sans avoir recours à un récepteur spécifique [203], [205], [213], [216]. Cependant, ces résultats ont été remis en question par l'équipe de B. Lebleu qui a constaté que la méthode utilisée pour étudier ces mécanismes d'internalisation, à savoir la microscopie confocale sur cellules fixées, pouvait être à l'origine d'artefacts dans la mesure où les peptides vecteurs accumulés à la membrane ont la possibilité de se redistribuer dans les cellules lors de l'étape de fixation [217]. C'est alors l'hypothèse d'une internalisation des CPPs par **endocytose** qui a été avancée. Depuis, de nombreuses études ont montré que différentes voies d'endocytose, mécanismes dépendants de l'ATP, étaient effectivement vraisemblablement responsables de la majeure partie de l'internalisation [218]–[220]. Toutefois, le processus de translocation directe est soutenu par des études biophysiques utilisant par exemple des vésicules unilamellaires [221], [222] et constituerait la voie d'internalisation principale pour les CPPs amphipathiques présents à de fortes concentrations dans le milieu extracellulaire [223].

De plus, la réalisation d'expériences d'internalisation à 4 °C, (température à laquelle les mécanismes dépendants de l'ATP sont inhibés) et à 37 °C pour différents peptides vecteurs a permis de montrer que les deux types de mécanismes étaient utilisés pour l'internalisation. La **contribution des deux mécanismes** dépend fortement de la séquence peptidique, de la concentration extracellulaire en peptide et de la présence d'éléments chargés négativement à la membrane tels que les glycosaminoglycanes (GAGs) [224].

Dans cette partie, nous souhaitons présenter quelques bases pour la compréhension des mécanismes d'internalisation des CPPs. Cependant étant donné l'extraordinaire complexité du sujet et par souci de concision, nous ne détaillerons pas les nombreuses hypothèses formulées pour expliquer la capacité des différents CPPs à s'internaliser par différents mécanismes en fonction de leur séquence, leur concentration, leur capacité à se structurer, etc... Nous retiendrons donc surtout les critères essentiels à l'interaction avec la membrane plasmique et les principaux mécanismes d'internalisation envisagés. Ensuite, nous aborderons la question cruciale des méthodes d'étude de l'internalisation des CPPs et verrons quels biais elles sont susceptibles de causer dans l'interprétation des résultats.

### **1. Description des mécanismes proposés pour l'internalisation des CPPs**

#### **a. Interaction avec la membrane plasmique**

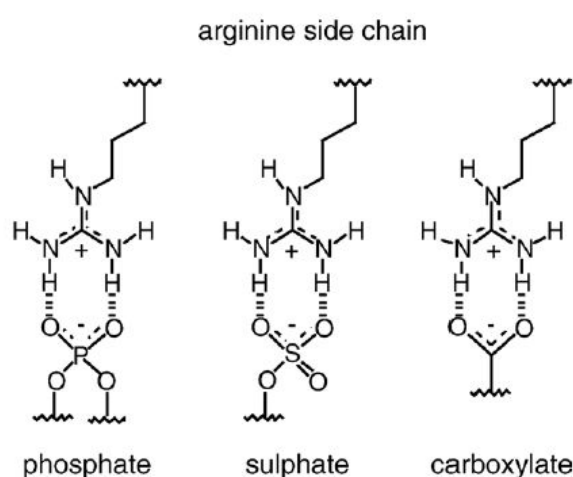
Quel que soit le mécanisme utilisé par les CPPs pour entrer dans les cellules, l'interaction avec la membrane plasmique est généralement considérée comme étant la première étape du processus d'internalisation.

La bicouche lipidique de la membrane plasmique délimite l'intérieur et l'extérieur d'une cellule. Cependant les molécules qui s'approchent d'une cellule vont d'abord rencontrer une couche dense appelée glycocalix et fortement chargée négativement. Ainsi, la membrane se présente comme un écran de charges négatives se plaçant devant une bicouche hydrophobe et ce sont à la fois des interactions électrostatiques et hydrophobes qui vont être responsables de l'affinité des CPPs avec la membrane des cellules.

### Rôle joué par les charges positives des CPPs :

Ce glycocalix comme son nom l'indique est constitué de glycoconjugués qui, en fonction de la nature de la partie sur laquelle se fixent les sucres et de la nature de ces sucres, sont répartis en quatre catégories : les glycoprotéines, les protéoglycanes, les glycophosphatidylinositol et les glycosphingolipides. Nous ne rentrerons pas dans les détails complexes de la nomenclature des sucres.

Le **glycocalix** confère à la surface de la cellule une forte charge négative et forme le premier obstacle aux molécules et aux organismes exogènes. Ainsi logiquement **la première interaction entre les CPPs chargés positivement et la membrane plasmique serait de nature électrostatique**, faisant intervenir les groupements négativement chargés de différents types de glycoconjugués. Ces interactions charge-charge qui ont une portée importante vont permettre d'attirer des CPPs vers la surface des cellules. Il a été montré que les résidus arginine des CPPs étaient plus favorables au processus d'internalisation que les résidus lysine [212]. Cette observation s'explique par la formation de complexes non-covalents « bidentates », c'est-à-dire impliquant deux liaisons hydrogènes entre les groupements guanidinium des arginines et les groupements phosphates, sulfates et carboxylates présents à la surface des cellules [225](cf. Figure 27).



**Figure 27 : Formation de liaisons hydrogènes divalentes entre les groupements guanidinium des chaînes latérales des arginines et des groupements phosphates, sulfates et carboxylates (Figure tirée de [226]).**

### **Rôle des résidus hydrophobes des CPPs :**

Les résidus hydrophobes, lorsqu'il y en a dans la séquence des CPPs, jouent un rôle majeur dans **l'interaction avec la bicouche lipidique de la membrane plasmique** et permettent d'améliorer l'internalisation des CPPs [227]. En effet, il a été observé que le remplacement de n'importe lequel des résidus hydrophobes du peptide pVec de séquence LLILL conduisait à une diminution de la quantité de peptide internalisée [228]. Il a été également montré que le nombre de résidus tryptophane était directement corrélé avec l'efficacité d'internalisation de différents CPPs [229].

### **Des partenaires d'interactions privilégiés ?**

Il est communément accepté que les CPPs ne rentrent pas dans les cellules *via* une interaction avec un récepteur spécifique. Cela est notamment lié au fait que leur internalisation peut se faire indépendamment de leur chiralité et sans phénomène de saturation [230]. Cependant des études récentes remettent en cause cette idée [231], [232].

Sans s'avancer à parler de récepteurs, il apparaît que les CPPs ont des partenaires d'interactions privilégiés à la membrane, c'est en particulier le cas des **glycosaminoglycanes (GAGs)**.

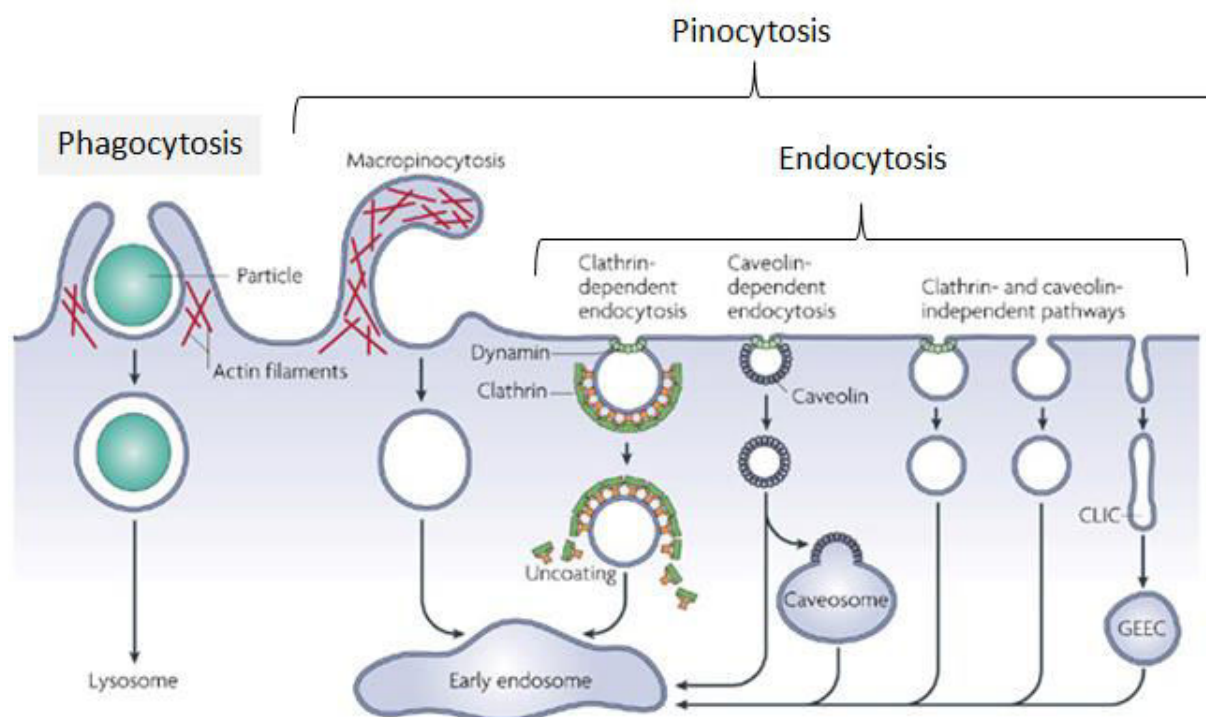
En effet, outre la formation d'interactions électrostatiques avec les CPPs, qui permet à ceux-ci d'approcher la membrane plasmique, les protéoglycanes, protéines mono- ou disubstituées par des polysaccharides à longues chaînes linéaires, polysulfatées et chargées négativement appelées GAGs, semblent jouer un rôle crucial dans les processus d'internalisation [233]. L'impact des GAGs sur l'internalisation a été mis en évidence en comparant la quantité de peptide internalisé pour des cellules déficientes ou non en GAGs. L'internalisation du peptide Tat est ainsi fortement diminuée lorsque les cellules sont traitées par des enzymes qui clivent les héparanes sulfates, une des classes de GAGs [234]. De même, l'entrée des différents CPPs est considérablement réduite dans les cellules CHO déficientes en GAGs [224]. Différentes hypothèses sont avancées pour justifier cette importance des GAGs, comme l'entrée des CPPs liés aux GAGs dans le cadre de leur cycle de recyclage à la membrane [235], [236] ou encore le déclenchement d'une aggrégation des GAGs par les CPPs [237].

Une fois le contact établi entre les CPPs et la membrane plasmique, les CPPs vont être internalisés vraisemblablement selon deux grandes familles de mécanismes : les mécanismes d'endocytose et les mécanismes de translocation directe, respectivement dépendants et indépendants d'un apport d'énergie.

#### **b. Endocytose**

L'endocytose est un processus naturel qui se produit dans tous les types cellulaires et qui correspond à une déformation de la membrane plasmique pour encapsuler dans une vésicule une molécule à internaliser. Le terme d'endocytose regroupe différents mécanismes qui nécessitent tous un apport d'énergie et qui ne sont donc efficaces qu'à température physiologique et en présence d'ATP.

Deux grands groupes de mécanismes peuvent être distingués : la **phagocytose** qui est réservée aux molécules de grandes tailles telles que les macrophages, et la **pinocytose** qui permet l'intégration de solutés variés de tailles plus modestes (*cf.* Figure 28).



**Figure 28 : Les différentes voies d'entrée par endocytose dans la cellule** (Figure tirée de [238]).

Le terme de pinocytose regroupe ensuite un certain nombre de mécanismes :

*La macropinocytose* : c'est un mécanisme qui implique des vésicules de relativement grande taille (0,5 à 2  $\mu\text{m}$ ) appelées macropinosomes. La formation de ces vésicules s'accompagne d'une polymérisation forte de l'actine à la surface des cellules qui entraîne une élévation de la membrane qui va se refermer sur elle-même en emprisonnant une grande quantité de fluide extracellulaire avec ce qu'il contient [239].

*L'endocytose dépendante de la clathrine* : c'est probablement la voie d'endocytose la mieux étudiée. Comme son nom l'indique, elle dépend de l'action de la clathrine membranaire, protéine qui est recrutée à la membrane et recouvre l'extérieur des vésicules formées ( $\approx 120 \text{ nm}$ ) en réalisant une structure de cage. Ce mécanisme d'endocytose implique également d'autres protéines telles que la dynamine [240].

*L'endocytose contrôlée par la cavéoline* : Ce mécanisme repose sur l'action de la protéine membranaire cavéoline qui va lier les molécules de cholestérol de la membrane pour former de petites vésicules ( $\approx 60 \text{ nm}$ ) appelées cavéoles et qui sont également riches en sphingolipides.

Enfin, d'autres voies d'endocytose, indépendantes des différents phénomènes cités ci-dessus et moins bien connues ont été observées. Elles sont rassemblées sous le terme vague d'endocytose indépendante de la clathrine et de la cavéoline.

Tous les types de pinocytoses connus ont été décrits comme permettant l'entrée de peptides vecteurs. Ainsi, l'endocytose dépendante de la clathrine a été proposée par Richard *et al.* [241] pour expliquer au moins partiellement l'internalisation du peptide Tat. En effet, une inhibition spécifique



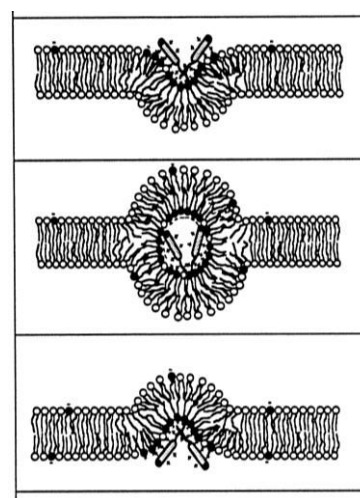
de cette voie a abouti à une diminution significative de la quantité de peptide Tat internalisé. De leur côté, Ferrari *et al.* ont observé une internalisation du peptide Tat porteur d'une étiquette GFP par la voie d'endocytose dépendante de la cavéoline [242]. Finalement, la macropinocytose a souvent été suggérée pour expliquer l'internalisation de Tat ou de peptides synthétiques riches en arginine [219], [220].

### c. Translocation directe :

Pour la majorité des CPPs, la quantité de peptide internalisée dans les cellules est diminuée à 4 °C ou lorsque des inhibiteurs d'endocytose sont utilisés [224], [243], [244]. Cependant le fait que cette capacité d'internalisation ne soit pas complètement supprimée indique la contribution de mécanismes d'internalisation indépendants de l'énergie rassemblés sous le terme de translocation directe.

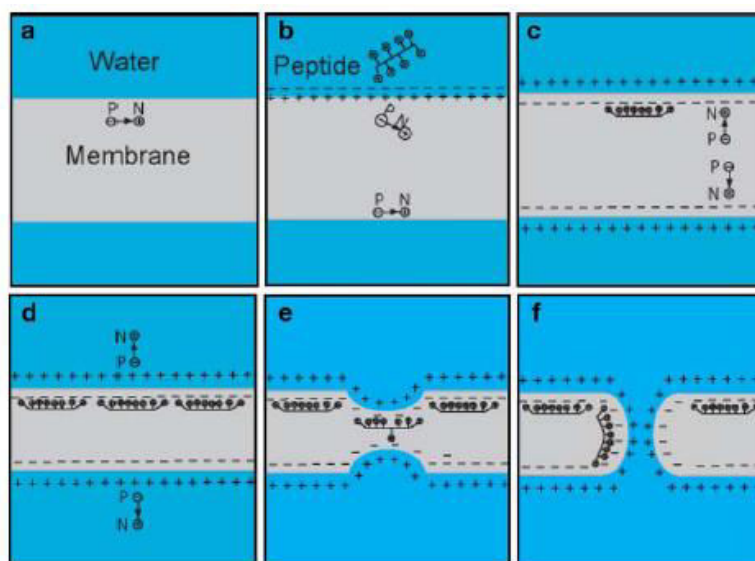
Trois principaux mécanismes de translocation directe ont été proposés.

*La formation de micelles inverses* : C'est le mécanisme qui a été proposé pour expliquer la translocation directe de la pénétratine [216] (*cf.* Figure 29). Dans ce modèle, les résidus basiques des CPPs interagissent avec les phospholipides négativement chargés de la membrane et cela induit une déstabilisation de la bicouche et une invagination de la membrane [245]. Dans le même temps, les lipides voisins vont se réorganiser, conduisant à la formation de micelles qui vont encapsuler le CPP [246].



**Figure 29 : Modèle proposé pour expliquer la translocation directe des CPPs par formation de micelles inverses** (Figure tirée de [216]).

*La formation de pores* : Des simulations de dynamique moléculaire et des expériences d'électrophysiologie suggèrent la formation de pores toroïdaux pour permettre la diffusion passive de Tat et des peptides riches en arginine à travers la membrane [247]. La formation de pores serait causée par l'accumulation de peptides dans le feuillet externe de la membrane, elle-même due à l'interaction entre les groupements positivement chargés des peptides et les groupements phosphates des lipides. Cette accumulation de peptides conduirait à un amincissement de la bicouche et à l'attraction entre les groupements des chaînes latérales des arginines et des lysines et les groupements phosphates de la couche interne, ce qui causerait la formation d'un pore transitoire (*cf.* Figure 30).



**Figure 30 : Schéma des différentes étapes de translocation directe par formation de pores.** Le vecteur délimité par les lettres P et N indique l'existence d'un moment dipolaire et donne son orientation. Un changement d'orientation de ce dipôle est causé par la présence de CPPs chargés positivement au niveau de la membrane externe. (Figure tirée de [247])

*Perméabilisation par électroporation* : Ce mécanisme a été décrit par Blinder et Lindblom et suggère qu'au-delà d'une certaine concentration en peptide, la distribution asymétrique des charges des deux côtés de la membrane conduirait à l'établissement d'une tension électrique transmembranaire. La membrane serait alors perméabilisée de manière semblable à de l'électroporation par la création de défauts ponctuels permettant la traversée de la membrane par les peptides de façon à rétablir un potentiel nul [221].

## 2. Méthodes d'études de la capacité d'internalisation des CPPs dans les cellules

Différentes méthodes ont été utilisées pour étudier les mécanismes d'internalisation, localiser et quantifier les CPPs dans les cellules. Nous allons brièvement présenter les principales méthodes, leurs applications et les précautions à prendre pour les mettre en œuvre.

### a. Les méthodes basées sur l'utilisation de la fluorescence :

Les méthodes les plus couramment utilisées reposent sur l'émission d'un signal de fluorescence pour visualiser ou quantifier le peptide internalisé. Une fois que le CPP a été couplé à un fluorophore, il est possible de le quantifier en fluorimétrie ou en cytométrie en flux [248] ou de le localiser en microscopie électronique [249]. Les expériences de microscopie de fluorescence visant à localiser le peptide sont majoritairement réalisées sur cellules vivantes de manière à limiter les redistributions artéfactuelles du peptide au moment de la fixation des cellules déjà évoquée précédemment [217].

Cependant l'introduction d'une sonde fluorescence sur une courte séquence peptidique tel qu'un CPP n'est pas anodine. En effet, les fluorophores sont généralement de grandes molécules très hydrophobes et il a été montré que le marquage d'un CPP avec un fluorophore pouvait modifier significativement son interaction avec la membrane plasmique [250], augmenter sa capacité

d'internalisation ainsi que sa distribution dans la cellule [251], [252]. De plus, il a été montré que les fluorophores étaient susceptibles de modifier la flexibilité et la conformation des peptides [253].

D'autres éléments peuvent poser problèmes pour la quantification des CPPs internalisés en fluorimétrie ou en cytométrie en flux. Tout d'abord, il faut s'assurer de la **stabilité de la liaison entre le CPP et le fluorophore**, pour que la fluorescence détectée corresponde au peptide internalisé et non à une étiquette fluorescente seule. Ensuite, un phénomène de **quenching de la fluorescence** peut se produire s'il y a une trop forte accumulation dans un compartiment cellulaire ou en présence de partenaires d'interactions [254]. Finalement, il faut garder à l'esprit que les CPPs ont une forte affinité avec la membrane des cellules et que des lavages répétés peuvent ne pas être suffisants pour éliminer le **peptide dit « membranaire »**. Cela est d'autant plus important que la quantité de peptide membranaire représente, la plupart du temps, une quantité bien supérieure à celle de peptide internalisé [255]. Pour éliminer la part de signal liée à la présence de peptide membranaire, il y a deux possibilités : (1) réaliser une étape de digestion du peptide à la membrane afin de le libérer dans le milieu extracellulaire, (2) utiliser un extincteur de fluorescence incapable de passer la membrane plasmique [255].

### **b. Une méthode de quantification par spectrométrie de masse**

Pour pallier les nombreux inconvénients de ces méthodes de quantification basées sur la fluorescence, une technique de quantification des CPPs internalisés par spectrométrie de masse a été mise au point au laboratoire [256]. Cette technique, à laquelle nous avons fait appel dans le cadre de ces travaux de thèse et qui sera plus amplement détaillée dans le chapitre 5, repose sur l'utilisation de peptides vecteurs biotinylés sous forme deutérée et non-deutérée. Tandis qu'une des deux versions du peptide vecteur est internalisée dans des cellules, l'autre version va être introduite en quantité connue au moment de la lyse des cellules (soit après leur internalisation) et servir d'étalon pour la quantification. Les peptides vecteurs biotinylés sont récupérés sur billes magnétiques de streptavidine et analysés en spectrométrie de masse MALDI-TOF. De la même manière que pour les expériences basées sur la fluorescence, il est nécessaire de s'affranchir de la part de peptide membranaire en réalisant une digestion enzymatique de ces derniers avant de lyser les cellules. La biotine est un groupement beaucoup moins volumineux et hydrophobe que les fluorophores et il a été montré qu'elle n'affectait pas l'interaction avec les membranes modèles [257]. Cette méthode constitue donc une alternative, plus robuste, à la quantification par fluorescence.

### **c. Les méthodes basées sur une activité biologique**

Dans la mesure où les CPPs ont pour application principale la vectorisation de molécules actives, il est tout à fait pertinent de vouloir utiliser des méthodes basées sur le suivi d'une activité biologique. Une expérience a été développée sur ce principe pour la vectorisation de protéines de choc thermique par des CPPs [258]. D'autre part, l'internalisation du peptide Tat a été étudiée et sa capacité à sortir des endosomes après son entrée par macropinocytose a été améliorée *via* sa conjugaison à la Cre recombinase [259], enzyme localisée dans le noyau.

Cependant si ces méthodes peuvent être bien adaptées à certains systèmes CPP- cargo, elles ne sont pas idéales pour étudier le devenir intracellulaire du CPP puisque celui-ci peut-être influencé par le

cargo, comme nous allons le voir dans la partie suivante. De plus, il est possible que le cargo ne soit pas convenablement modifié par sa cible ou que sa transformation soit réversible ce qui va empêcher de rationaliser correctement les résultats et compliquer l'interprétation en termes de mécanismes d'internalisation.

### III. Seulement des peptides vecteurs ?

#### 1. Les domaines d'applications des CPPs

La plupart du temps, les CPPs sont destinés à la vectorisation dans les cellules de molécules biologiquement actives appelées cargos (et non cargaisons contrairement à ce qu'on pourrait penser) et qui peuvent être des peptides, des protéines, des ARNs ou encore, par exemple, des nanoparticules [260]–[262].

Les applications potentielles sont donc en majeure partie dans le domaine médical avec la vectorisation de médicaments, la thérapie génique [263] ou la thérapie ciblée contre le cancer [264], mais des développements biotechnologiques comme la vectorisation de sondes ou de biocapteurs pour la recherche en biologie cellulaire peuvent également être envisagés [265], [266]. Les cargos sont généralement covalamment liés aux CPPs par un pont disulfure qui sera réduit lors de l'arrivée du complexe dans le cytosol, milieu réducteur. Des protéines de fusion exprimées avec des CPPs peuvent également être utilisées [267]. Finalement, il a été montré que l'internalisation de cargos non-covalamment liés aux CPPs était possible [268].

#### 2. L'influence des cargos sur le devenir intracellulaire des conjugués

Une étude poussée menée avec trois CPPs (Tat, Tp10 et pénétratine) porteurs de différents cargos a mis en évidence l'influence du cargo à la fois sur l'efficacité de l'internalisation et sur la cytotoxicité des conjugués [269]. Tünnemann *et al.* se sont, eux, intéressés à **l'impact du cargo sur le mode d'internalisation** : ils ont comparé la vectorisation par le peptide Tat de cargos de tailles différentes et ont montré que les conjugués de grandes tailles (association avec des quantum dots (nanoparticules de semi-conducteurs avec des propriétés quantiques utilisées en imagerie cellulaire) ou protéines de fusion) étaient essentiellement internalisés par endocytose tandis que la part de translocation directe était beaucoup plus importante dans le cas de la vectorisation de petits peptides [270].

Le mode d'internalisation des conjugués est un élément déterminant. En effet, pour qu'un cargo puisse jouer un rôle actif dans une cellule, il faut qu'elle puisse atteindre le compartiment cible. Si le conjugué rentre par translocation directe, il va se retrouver directement dans le cytosol. En revanche, un conjugué internalisé par **endocytose sera encapsulé dans des endosomes** dans lesquels il peut rester piéger. Ainsi, si le cargo à vectoriser est de taille importante, il sera important de réfléchir à des façons de promouvoir la fuite des endosomes. Différentes techniques existent telles que l'association à des polymères comme le polyéthylène imine (PEI) [271] ou à des séquences d'acides aminés particulières. Ces techniques reposent sur une perturbation de la membrane endosomale [259].

Atteindre le cytosol peut donc déjà être difficile pour un conjugué, et cela se corse encore si le cargo doit atteindre un autre compartiment intracellulaire tel que le noyau ou les mitochondries. Dans le cas d'un ciblage intracellulaire précis, il faudra intégrer au conjugué des séquences spécifiques d'adressage comme, par exemple, un domaine dérivé du signal de localisation nucléaire (NLS) de l'antigène SV40 pour un ciblage du noyau [210].

Par conséquent, il est absolument nécessaire de disposer d'une bonne connaissance des mécanismes d'internalisation, de libération dans le cytosol ou de ciblage intracellulaire de façon à concevoir des conjugués CPP-cargo qui seront capables d'atteindre leur cible.

### 3. Des peptides vecteurs avec une activité biologique intrinsèque

Les peptides vecteurs sont généralement considérés comme inertes, permettant uniquement le passage de composés bioactifs au travers de la membrane plasmique des cellules ou éventuellement également leur transport jusqu'à un compartiment intracellulaire cible.

Cependant, aujourd'hui, après près de trois décennies d'études, il apparaît qu'une partie des CPPs ne sont pas de simples vecteurs, utiles pour moduler la pharmacocinétique de composés bioactifs, mais qu'ils présentent également une **activité biologique intrinsèque**. Ainsi, il a notamment été montré qu'un certain nombre d'entre eux ont des propriétés antimicrobiennes. C'est le cas par exemple de Tat ou de la pénétratine [272].

Le terme de bioportide a récemment été introduit pour désigner ces CPPs ayant une activité biologique intrinsèque [1].

En 2011, une revue de W. Verdurmen et R. Brock rassemble différents exemples de CPPs bioactifs ou bioportides et de leurs effets observés [231]. Il en ressort que les effets observés sont très variés et qu'ils ne peuvent pas être nécessairement directement reliés au processus d'internalisation. Les auteurs insistent sur la **nécessité de mieux comprendre les bases moléculaires de ces effets** et cela notamment pour être en mesure de les imputer à des enchaînements d'acides aminés qui pourront ensuite être repris pour le design rationnel de nouvelles molécules bioactives qui ne possèderaient que les propriétés délibérément introduites. Parmi les exemples cités figure celui de l'étude de l'effet biologique intrinsèque du peptide (R/W)<sub>9</sub> menée au laboratoire (UMR 7213, UMR 7203) lors des travaux de thèse de Delaroche D. [2]. Ces travaux ont permis de mettre en évidence la capacité de remodelage du cytosquelette d'actine de fibroblastes tumoraux par le CPP (R/W)<sub>9</sub>, ainsi qu'une diminution de la motilité et de la capacité à croître en indépendance d'ancrage après quelques heures d'incubation avec ce peptide.

Au cours de ces trois années de thèse, j'ai cherché à obtenir une meilleure compréhension des bases moléculaires de l'effet observé du peptide (R/W)<sub>9</sub> sur le phénotype des cellules tumorales EF. En particulier, nous avons cherché à trouver des partenaires d'interaction de ce peptide tout au long du chemin parcouru par (R/W)<sub>9</sub> depuis l'extérieur vers l'intérieur des cellules, au cours de son internalisation dans les cellules EF et nous sommes intéressés à son effet global sur l'expression des protéines.



# Présentation du projet de recherche

## I. Contexte de l'étude

Ce travail de thèse est le prolongement direct d'une étude menée en grande partie au cours de la thèse de Diane Delaroche intitulée « *Effets de peptides vecteurs sur le cytosquelette d'actine après internalisation dans des cellules tumorales* », réalisée sous la direction de Sandrine Sagan et soutenue le 15 novembre 2007.

Ainsi pour préciser clairement le contexte de mon projet de thèse, il est essentiel que je rappelle les principaux résultats obtenus au cours de ces précédents travaux [2].

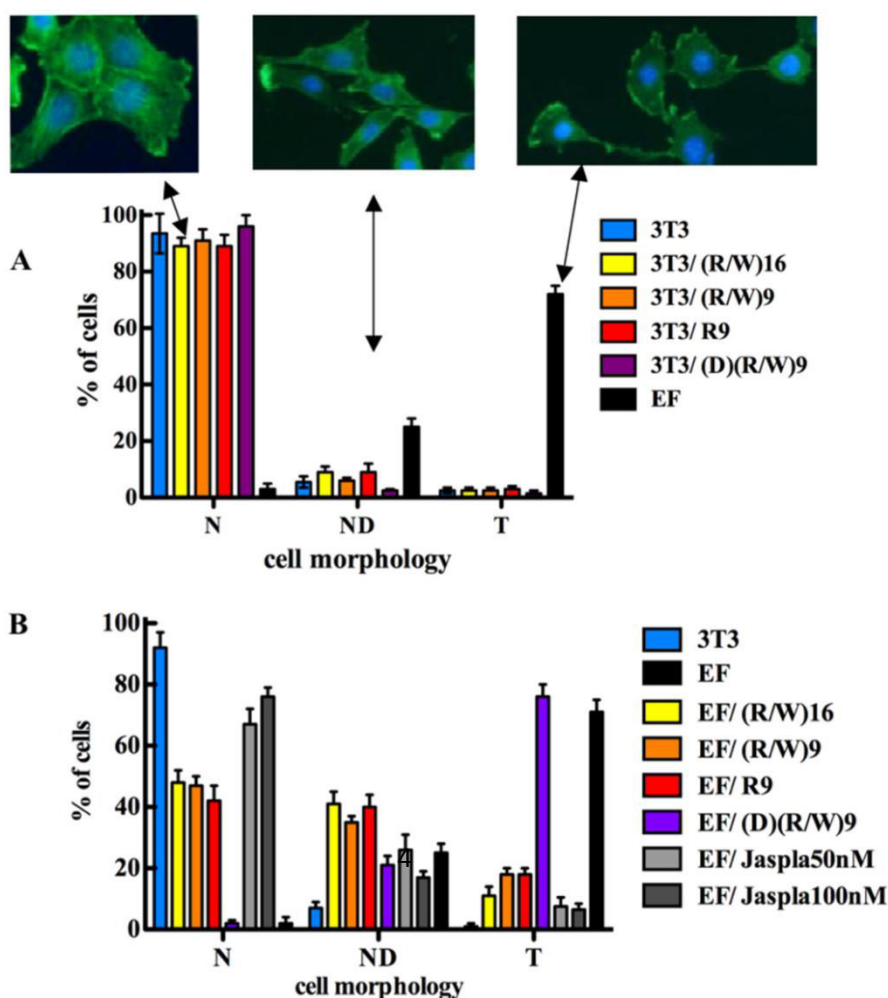
Ces travaux sont au départ basés sur l'observation suivante : certains CPPs ont un effet sur la polymérisation des filaments d'actine. Cet effet a été observé *in vitro* pour les peptides (R/W)<sub>16</sub> de séquence RRWRRWRRWRRWRR, (R/W)<sub>9</sub> de séquence RRWRRWRR et R<sub>9</sub> de séquence RRRRRRRR grâce à la méthode de mesure de la polymérisation de l'actine par anisotropie de fluorescence mise au point par Spitz J.A. *et al.* [273]. Lors de ce screening l'effet des peptides pénétratine (ERQIKVFQNRRMKWKK) et (R/L)<sub>16</sub> (RRLRLLRRLRLRR) a également été évalué mais aucune influence significative sur la vitesse de polymérisation de l'actine n'a pu être mise en évidence.

Suite à ce premier test, l'effet des peptides (R/W)<sub>16</sub>, (R/W)<sub>9</sub> et R<sub>9</sub> sur la lignée cellulaire EF a été évalué. La lignée cellulaire EF est issue de la transformation tumorale de fibroblastes murins de type NIH-3T3 par la protéine de fusion oncogénique EWS-FLI1. Cette lignée cellulaire est un modèle du sarcome d'Ewing, cancer des os et tissus conjonctif déjà décrit dans la partie III.2 du chapitre 1. Ces cellules tumorales EF présentent un cytosquelette d'actine très désorganisé dépourvu de fibres de stress (association de filaments d'actine en faisceaux, cf. chap 1, partie II. 3. b.) et constituent ainsi un modèle intéressant pour évaluer un effet sur la polymérisation de l'actine.

Tout d'abord, il a été montré par des expériences d'immunofluorescence que les peptides (R/W)<sub>16</sub>, (R/W)<sub>9</sub> et R<sub>9</sub> induisent un **changement de morphologie des cellules EF**. Tandis que les cellules 3T3, non tumorales, présentent une forme d'étoile liée à l'existence d'un important réseau de fibres de stress et à un ancrage solide à leur support, les cellules EF, tumorales, ont une forme arrondie en raison de la perte de leurs points d'adhésion focaux et de leurs fibres de stress. Après une incubation comprise entre 18 et 24 h avec 10 µM des différents peptides, un retour à une morphologie typique des cellules 3T3 accompagnée de la réapparition de fibres de stress ont été observés de manière reproductible pour une part importante des cellules. Le changement de morphologie des cellules EF et des cellules 3T3 en présence des peptides (R/W)<sub>16</sub>, (R/W)<sub>9</sub> et R<sub>9</sub> a été suivi de manière quantitative en définissant différentes classes de cellules en fonctions de leur morphologie : une classe de cellules semblables aux non-tumorales (N), une classe de cellules avec une morphologie non modifiée par le traitement soit des cellules tumorales (T) et une classe pour les cellules ne pouvant être facilement attribuées à l'une ou l'autre des deux catégories précédentes (ND). Le peptide (R/W)<sub>9</sub> (D) composé uniquement d'acides aminés de configuration D a également été testé mais, contrairement à la version de (R/W)<sub>9</sub> composée d'acides aminés naturels de configuration L, aucun remodelage du

cytosquelette d'actine n'a été observé. Cela indique qu'une reconnaissance chirale serait impliquée dans le mécanisme de remodelage du cytosquelette d'actine. L'effet des CPPs (R/W)<sub>16</sub>, (R/W)<sub>9</sub> et R<sub>9</sub> a été comparé avec celui de la jasplakinolide, petite molécule d'environ 700 Da connue pour entraîner une polymérisation de l'actine [274]. L'effet des CPPs sur le remodelage du cytosquelette d'actine est un peu moins marqué que celui de la jasplakinolide : en moyenne 45 % des cellules retrouvent une morphologie non-tumorale après traitements avec les différents CPPs contre 70 à 80 % pour la jasplakinolide. Ainsi les populations de cellules obtenues à l'issue du traitement avec les peptides sont hétérogènes.

Les résultats de ces expériences sont présentés dans le graphe ci-dessous (Figure 31).



**Figure 31 : Analyse quantitative des effets de différents peptides vecteurs sur la morphologie des cellules 3T3 et EF.** Dans trois expériences indépendantes plus de 500 cellules 3T3 (A) ou EF (B) ont été examinées et classées en fonction de leur morphologie dans les catégories « non-tumorales » (N), « tumorales » (T) ou « indéterminées » (ND) en l'absence ou en présence de 10  $\mu$ M des peptides (R/W)<sub>16</sub>, (R/W)<sub>9</sub>, R<sub>9</sub> et (R/W)<sub>9</sub>(D) et avec 50 ou 100 nM de jasplakinolide (Figure issue de [2]).

Ensuite, la motilité des cellules EF a été mesurée pendant 18 h par vidéomicroscopie en présence du peptide (R/W)<sub>16</sub> (10  $\mu$ M). Cette expérience a permis de montrer que le (R/W)<sub>16</sub> induisait une **diminution de la motilité des cellules EF** jusqu'à des valeurs comparables à celles des cellules 3T3.



Finalement, la **capacité des cellules EF à croître en indépendance d'ancrage** à un support a été évaluée en l'absence et en présence des peptides (R/W)<sub>9</sub>, (R/W)<sub>16</sub> et R<sub>9</sub> en se basant sur le dénombrement des clones obtenus en milieu semi-solide. Tandis qu'aucun clone n'a été obtenu pour les cellules 3T3, une diminution significative par rapport aux nombres de clones obtenus pour les cellules EF de référence a été mise en évidence pour les cellules incubées en présence des peptides vecteurs avec une diminution de l'ordre de 30 % avec (R/W)<sub>16</sub> et (R/W)<sub>9</sub> et de 70 % pour R<sub>9</sub> à des concentrations pour lesquelles ces derniers ne présentent pas de cytotoxicité.

Ces différents éléments vont dans le sens d'une « **réversion** » du **phénotype tumoral des cellules EF par les peptides vecteurs**.

Etant donné l'effet préalablement observé *in vitro* sur la polymérisation de l'actine par anisotropie de fluorescence, la première hypothèse avancée pour expliquer les résultats obtenus sur les cellules EF a été celle d'une interaction directe avec l'actine. Afin de confirmer cette interaction directe avec l'actine et de mieux la caractériser, de nouvelles expériences de calorimétrie par titrage isotherme (ITC) et de RMN ont été réalisées. Les expériences de calorimétrie ont montré que les peptides (R/W)<sub>9</sub>, (R/W)<sub>16</sub> et R<sub>9</sub> interagissaient avec l'actine G avec des affinités respectives de 0,4, 10 et supérieure à 10  $\mu$ M (valeurs d'affinités relativement modestes en biologie). Les expériences de RMN menées avec (R/W)<sub>9</sub> et (R/W)<sub>16</sub> ont confirmé cette interaction *in vitro* et des expériences de compétition ont montré la capacité de (R/W)<sub>16</sub> à déplacer la thymosine  $\beta$ 4 fixée à l'actine G.

Malgré ces différents résultats, il n'est pas évident de trouver une interprétation cohérente quant à l'action des peptides (R/W)<sub>9</sub>, (R/W)<sub>16</sub> et R<sub>9</sub> sur un ou plusieurs partenaires intracellulaires de l'actine, ni de proposer une voie de signalisation pouvant conduire à cette « réversion » du phénotype tumoral.

## II. Objectif et structure de ce projet de thèse

A partir de ces conclusions, l'objectif de la première partie de mon travail de thèse a été de confirmer l'interaction des peptides (R/W)<sub>9</sub> et (R/W)<sub>16</sub> avec l'actine *in vitro* mais surtout de caractériser la (ou les) zone(s) d'interaction par des expériences de cross-linking chimique couplées à la spectrométrie de masse. Cette étude qui a fait l'objet d'une publication est décrite dans le chapitre 4.

A l'issue de ces expériences réalisées *in vitro*, nous avons souhaité nous rapprocher de la réalité de la cellule entière et essayer d'identifier sans *a priori* de nouveaux partenaires d'interaction et/ou de nouvelles pistes de voies de signalisation pour comprendre l'effet observé de ces peptides sur le phénotype des cellules EF. Nous avons choisi de nous focaliser sur le peptide (R/W)<sub>9</sub> car, des trois peptides vecteurs présentant une activité biologique, nous avons jugé qu'il était le CPP le plus prometteur (cytotoxicité plus faible que (R/W)<sub>16</sub> et davantage d'informations sur son effet biologique que pour R<sub>9</sub>).

L'objectif global de ce projet de thèse est le suivant :

**Comprendre l'effet moléculaire du peptide vecteur (R/W)<sub>9</sub> sur le phénotype de cellules modèles du sarcome d'Ewing.**

Pour progresser vers cet objectif, nous nous sommes appuyés sur deux stratégies complémentaires : une **stratégie ciblée** et une **stratégie de protéomique différentielle globale**.

Le principe de la stratégie ciblée est le suivant : capturer et identifier des partenaires d'interaction du peptide pour approcher son mécanisme d'action. Dans la stratégie ciblée nous allons distinguer deux types d'expériences : les expériences reposant sur le **photocross-linking** réalisées majoritairement *in cellulo* qui sont l'objet du chapitre 5 et les expériences de **pull-down** qui sont décrites dans le chapitre 6.

Le principe de la stratégie différentielle est, quant à lui, d'évaluer l'influence, l'impact de la présence du peptide (R/W)<sub>9</sub> dans le milieu cellulaire en comparant le protéome des cellules EF incubées avec le peptide par rapport à celui de cellules identiques mais non soumises à l'action de (R/W)<sub>9</sub>. Cette stratégie de protéomique différentielle globale s'est appuyée sur la technique de **protéomique quantitative SILAC** et est détaillée dans le chapitre 7.

## Chapitre 4 : Cross-linking chimique couplé à la spectrométrie de masse pour étudier l'interaction des CPPs (R/W)<sub>9</sub> et (R/W)<sub>16</sub> avec l'actine.

### I. Objectifs de l'article 1

L'étude menée au laboratoire sur l'effet des peptides vecteurs (R/W)<sub>9</sub> et (R/W)<sub>16</sub> sur les cellules tumorales EF a permis de mettre en évidence leur activité de remodelage de l'actine. De plus, différentes expériences réalisées *in vitro* avec de l'actine G ont indiqué la possibilité d'une interaction directe avec l'actine pour expliquer l'effet des peptides sur le phénotype des cellules EF [2]. Nous avons voulu aller plus loin dans la caractérisation de l'interaction des peptides (R/W)<sub>9</sub> et (R/W)<sub>16</sub> avec l'actine G et avons pour cela choisi de mettre au point une méthode de cross-linking chimique couplée à la spectrométrie de masse. Ce travail d'exploration réalisé *in vitro* en cross-linking chimique nous a permis de mettre au point sur un système simple (une protéine et un peptide vecteur) les conditions optimales pour réaliser une réaction de cross-linking. Par ailleurs, ceci nous a permis de mettre au point les conditions de purification d'affinité des complexes cross-linkés (biotine/streptavidine) et leur analyse par spectrométrie de masse MALDI-TOF/TOF et nanoLC-ESI-LTQ/Orbitrap. Ce travail s'est révélé indispensable pour la suite de notre étude réalisée dans des milieux beaucoup plus complexes.

De manière à obtenir un maximum d'informations, nous avons intégré dans notre méthode deux stratégies complémentaires : la première stratégie repose sur **l'analyse des complexes entiers** *via* une analyse en spectrométrie de masse MALDI-TOF en mode linéaire des milieux de cross-linking, la deuxième stratégie, qui est qualifiée de stratégie bottom-up, consiste en l'analyse par spectrométrie de masse en tandem des **espèces cross-linkées obtenues après digestion enzymatique** (trypsine) des complexes. La première technique vise à obtenir la composition des complexes entiers, la deuxième a pour objectif de préciser les zones d'interactions des CPPs avec les protéines.

En parallèle de l'étude réalisée avec l'actine, nous avons travaillé avec l'albumine, une protéine modèle, qui nous a permis de mieux appréhender les complexes réalisés avec l'actine. Cependant, cette protéine n'a pas été choisie au hasard, en effet, Futaki *et al.* ont montré que des protéines du sérum telles que l'albumine interagissaient avec les CPPs riches en arginines et pouvaient modifier leur internalisation [275]. Ainsi nous avons cherché à voir si cette interaction existait aussi dans le cas des CPPs (R/W)<sub>9</sub> et (R/W)<sub>16</sub>.

### II. Principaux résultats

#### 1. Compositions des complexes

En MALDI-TOF, le caractère acide de la matrice ne permet pas de préserver les interactions non covalentes. Par contre, cette méthode est parfaitement adaptée pour l'étude de complexes entiers stabilisés par cross-linking chimique.

Pour décrire de manière quantitative les spectres MALDI-TOF obtenus pour les complexes entiers cross-linkés, nous avons développé un outil appelé SIMUL-XL nous permettant d'obtenir une estimation précise du nombre moyen de cross-linkers et de CPPs entrant dans la composition des complexes. Ce programme SIMUL-XL, développé par Gérard Bolbach (Visual basic), nous a, par exemple, permis de déterminer que pour un mélange 1:1 (R/W)<sub>9</sub>-albumine (10 µM-10 µM), le nombre moyen de CPP covalamment fixé est  $\lambda = 0,95$  avec un nombre moyen de cross-linkers  $\langle N=49 \rangle$  (écart-type  $\sigma = 5$ ) (cf. figure 2 de l'article). Pour un mélange 1 : 1 (R/W)<sub>9</sub>-actine (10 µM-10 µM), le nombre moyen de CPP covalamment fixé est  $\lambda = 0,3$  avec un nombre moyen de cross-linkers  $\langle N = 8 \rangle$  (écart type  $\sigma = 2$ ) indiquant un rendement de cross-linking plus faible dans le cas de ce deuxième système.

## 2. Zones d'interaction

L'identification des zones d'interaction à partir des spectres MS et MS/MS des espèces cross-linkées a pu être réalisée de manière automatisée avec l'utilisation du logiciel Xlink-Identifier développé par nos collaborateurs [150] et adapté pour notre système.

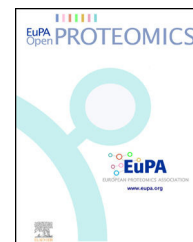
Les mêmes zones d'interactions ont été identifiées pour les CPPs (R/W)<sub>9</sub> et (R/W)<sub>16</sub> en complexes avec les protéines actine et albumine.

Pour les systèmes (R/W)<sub>9</sub> ou (R/W)<sub>16</sub>-actine 4 zones d'interactions de la protéine ont été mises en évidence : La zone [35–57] comprenant les lysines cross-linkées K45 et K56, la zone [91–111] comprenant la lysine cross-linkée K108, la zone [202–210] comprenant la lysine cross-linkée K208 et la zone [308–333] comprenant les lysines cross-linkées K310, K321 et K323. On peut s'étonner de ne pas trouver une plus grande spécificité d'interaction avec un ou deux sites privilégiés. Il se peut que la nature à la fois chargée positivement et hydrophobe confère à la fois la possibilité aux peptides (R/W)<sub>9</sub> et (R/W)<sub>16</sub> d'interagir de manière spécifique mais aussi de se lier via des interactions non spécifiques simplement du fait de leur composition chimique. Cependant, parmi ces quatre zones deux, [46–57] et [202–210], pourraient avoir un intérêt biologique puisqu'elles sont connues pour leur implication dans la modulation de la dynamique de l'actine [276]. De plus, la zone d'interaction [46–57] fait partie du domaine de liaison à l'actine de la thymosine  $\beta 4$  et pourrait par conséquent expliquer le déplacement de la thymosine  $\beta 4$  par (R/W)<sub>16</sub> observé lors des expériences de compétition en RMN [2].

Pour les systèmes (R/W)<sub>9</sub> ou (R/W)<sub>16</sub>-albumine cinq zones d'interaction ont été identifiées [24–44], [210–218], [233–248], [452–459] et [548–557] quel que soit le ratio protéine : CPP utilisé (1:1 ou 10:1) indiquant que parmi ces différents sites aucun n'est colonisé de manière préférentielle. L'albumine semble être capable de fixer un grand nombre de CPPs, dans des zones multiples et réparties sur sa surface exposée, sans spécificité apparente, ce qui va dans le sens du rôle de transport/protection que cette protéine exerce dans le sang.

Available online at [www.sciencedirect.com](http://www.sciencedirect.com)

ScienceDirect

journal homepage: <http://www.elsevier.com/locate/euprot>

# An integrated cross-linking-MS approach to investigate cell penetrating peptides interacting partners

S  verine Clavier<sup>a,b</sup>, Xiuxia Du<sup>c</sup>, Sandrine Sagan<sup>a</sup>, G  rard Bolbach<sup>a,b</sup>,  
Emmanuelle Sachon<sup>a,b,\*</sup>

<sup>a</sup> Laboratoire des Biomol  cules, UMR7203 UPMC-ENS-CNRS-INSERM ERL1157, cc182, 4 place Jussieu, 75252 Paris Cedex 05, France

<sup>b</sup> Plateforme de spectrom  trie de masse et prot  omique, IBPS-UPMC, cc41, 7-9 quai saint Bernard, 75252 Paris Cedex 05, France

<sup>c</sup> Department of Bioinformatics and Genomics, University of North Carolina at Charlotte, USA

## ARTICLE INFO

### Article history:

Received 25 November 2013

Received in revised form

7 March 2014

Accepted 11 March 2014

Available online 24 March 2014

### Keywords:

Chemical cross-linking  
Cell penetrating peptides  
Interacting partners  
Mass spectrometry  
Actin  
Albumin

## ABSTRACT

Cell penetrating peptides (CPPs) are attracting attention because of their ability to deliver biologically active molecules into cells. On their way they can interact with membrane and intracellular proteins. To fully understand and improve CPP efficiency as drug delivery tools, their partners need to be identified. To investigate CPP-protein complexes, chemical cross-linking coupled to mass spectrometry is a relevant method. With this aim, we developed an original approach based on two parallel strategies, an intact complex analysis and a bottom-up one, to have a global characterization of the cross-linked complexes composition as well as a detailed mapping of the interaction zones.

**Biological significance:** The robust and efficient cross-linking-MS workflow presented here can easily be adapted to any CPP-protein interacting system and could thus contribute to a better understanding of CPPs activity as cell-specific drug delivery tools. We validated the relevancy of this cross-linking-MS approach with two biologically active CPPs, (R/W)<sub>9</sub> and (R/W)<sub>16</sub>, and two interacting protein partners, actin and albumin, previously reported using isothermal titration calorimetry (ITC) and NMR. Cross-linking-MS results obtained on these previous studies allowed us to go further by providing a detailed mapping of the interaction zones. The identified interaction zones between actin and CPPs (R/W)<sub>9</sub> and (R/W)<sub>16</sub> are biologically meaningful. Two cross-linked zones [46–57] and [202–210] of actin are indeed involved in the modulation of its dynamics. Moreover, [46–57] domain has also been described as one interaction domain for thymosin   4 whose actin binding can be displaced by competition with (R/W)<sub>16</sub> (NMR experiments).

   2014 The Authors. Published by Elsevier B.V. on behalf of European Proteomics Association (EuPA). This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/3.0/>).

\* Corresponding author at: Laboratoire des Biomol  cules, UMR7203 UPMC-ENS-CNRS-INSERM ERL1157 Plateforme de spectrom  trie de masse et prot  omique, IBPS-UPMC, France. Tel.: +33 144273234.

E-mail address: [emmanuelle.sachon@upmc.fr](mailto:emmanuelle.sachon@upmc.fr) (E. Sachon).

<http://dx.doi.org/10.1016/j.euprot.2014.03.002>

2212-9685/   2014 The Authors. Published by Elsevier B.V. on behalf of European Proteomics Association (EuPA). This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/3.0/>).

## 1. Introduction

Twenty years ago, the discovery of peptides able to ubiquitously cross cellular membranes commonly named cell penetrating peptides (CPPs), with very limited toxicity, launched a novel field in molecular delivery based on these non-invasive vectors. Most CPPs are positively charged peptides though the presence of few anionic or hydrophobic CPPs was also demonstrated. After a decade of debate on the trafficking routes of CPPs to the heart of cells, it is now more or less accepted that these peptides use concomitantly different internalization pathways, including pinocytosis and direct membrane translocation processes [1]. CPPs are generally considered as biologically inert intracellular delivery tools. However, some CPPs have intrinsically biological activity and are part of a recently described class of CPPs baptized biopptide [2].

For instance, previous studies showed, that only the two CPPs (R/W)<sub>9</sub> and (R/W)<sub>16</sub> (RRWRRWR and RRWRRWRWRR respectively) among other CPPs tested, are able to remodel the actin cytoskeleton in oncogen transformed NIH3T3/EWS-Fli cells once these CPPs had crossed the plasma-membrane [3]. In order to explain the actin-remodeling activity of the two CPPs, the hypothesis of a direct interaction with actin was tested. (R/W)<sub>9</sub> and (R/W)<sub>16</sub> peptides were actually found to directly interact *in vitro* with G-actin by NMR and ITC experiments [3] ( $K_d \approx 10 \mu\text{M}$  and  $K_d = 0.4 \mu\text{M}$ , respectively). In addition, competitive binding experiments by NMR showed that (R/W)<sub>16</sub> was able to displace the actin sequestering protein thymosin  $\beta_4$  from G-actin [3].

It was also recently reported that arginine-rich CPPs interact with serum proteins like albumin, modifying their ability to internalize in cells [4]. Therefore our aim in this study was to analyze further the interaction of (R/W)<sub>9</sub> and (R/W)<sub>16</sub> with actin and albumin.

For this purpose, chemical cross-linking was chosen. Cross-linking reactions are conventionally based on the use of a bifunctional cross-linker, which is a carbon chain spacer bearing reactive sites at both ends that can either be identical (homobifunctional) or different (heterobifunctional). Reactive sites are mainly activated esters targeting either lysine residues (e.g. N-hydroxysuccinimide (NHS) ester) or cysteine residues (e.g. maleimide ester) although side reactions with tyrosine, threonine and serine have been reported [5].

By creating a covalent bond between two or more interacting partners, chemical cross-linking gives a snapshot of a molecule's environment and, combined with MS, constitutes a powerful tool to map protein-protein or peptide-protein interactions (distance constraints and interacting domains). However chemical cross-linking is often characterized by low reaction yields. In addition, a wide variety of cross-linked products are usually created. Therefore, cross-linkers bearing an affinity tag [6,7] allowing selective enrichment of the sample in cross-linked species have been developed. Cleavable cross-linkers [8,9], isotope labeled cross-linkers [10] or cross-linkers bearing a CHCA matrix moiety for MALDI analysis [11] improve detection and identification/characterization of the cross-linked species. Affinity purification using a

tagged cross-linker or interacting partners [12] and SDS-PAGE are the most commonly used off-line techniques but chromatographic methods like strong cation-exchange (SCX) [13,14] and size exclusion chromatography (SEC) [15] constitute promising approaches to enrich samples in cross-linked species.

Analysis of low abundant cross-linked peptides requires high sensitivity for their detection and high mass accuracy (mass error < 10 ppm) for their identification since the number of combination of two peptides is enormous. Moreover tandem MS of cross-linked peptides is required for the characterization of the interaction zones.

A number of tailored softwares have been developed to deal with the selective acquisition or interpretation of these MS/MS spectra. Among others, FINDX [16] has been designed to selectively fragment inter-protein cross-links by LC-MALDI-TOF/TOF using  $^{14}\text{N}/^{15}\text{N}$  mixed isotope strategy, xQuest [13] is dedicated to the search of isotopically tagged cross-linked peptides and CrossWork [17] or Xlink-Identifier [18] support label-free analyses of chemical cross-linking samples.

In this study, we set-up a general *in vitro* analytical workflow coupling cross-linking and mass spectrometry (cross-linking-MS), involving enrichment steps as well as manual or automated MS and MS/MS data processing to test potential interacting partners of any CPP sequence.

To validate our cross-linking-MS workflow, we studied the systems described above: (R/W)<sub>9</sub> or (R/W)<sub>16</sub> interacting with actin or albumin. For this purpose, CPP analogs suitable for chemical cross-linking experiments, were synthesized with the following sequences: Biot(O<sub>2</sub>)-G<sub>4</sub>-K-RRWRRWR-NH<sub>2</sub> and Biot(O<sub>2</sub>)-G<sub>4</sub>-K-RRWRRWRWRRWRWRR-NH<sub>2</sub>, respectively. The K residue added at the N-terminus of the peptides sequences allowed the cross-linking reaction, the biotin tag (Biot(O<sub>2</sub>)) was added for purification purpose and was separated from the biologically active motif by a four G residues spacer. This {Biot(O<sub>2</sub>)-G<sub>n</sub>-K-} group is easy to add at the N-terminus of peptides either manually or automatically during peptide synthesis whatever the peptide sequence.

The originality of our approach resides in the comprehensive study of the cross-linking reaction mixture from two angles: a global view of the interacting system with an intact complex analysis combined to the precise characterization of interacting zones by a bottom-up analysis. The intact complex analysis is based on the MALDI-TOF analysis in linear mode of the cross-linking reaction mixtures and on the modeling of the spectra obtained using the in-house SIMUL-XL program. The bottom-up analysis consists in the tryptic digestion of the cross-linking reaction mixtures followed by the affinity purification (biotin/streptavidin) of the biotinylated cross-linked peptides and their analysis by tandem MS (MALDI-TOF/TOF and/or nanoLC-ESI-Orbitrap). The MS/MS data are either manually interpreted with the help of GPMW software [19] (MALDI-TOF/TOF spectra) or automatically searched using Xlink-Identifier software [18] (nanoLC-ESI-MS/MS spectra).

For practical reasons the intact complex analysis was developed only with (R/W)<sub>9</sub> and both actin and albumin proteins. In contrast, the bottom-up analysis was performed using both (R/W)<sub>9</sub> and (R/W)<sub>16</sub> and both actin and albumin proteins.



## 2. Experimental

### 2.1. Materials

Bovine serum albumin and actin from rabbit muscle were purchased from Sigma (St. Louis, Missouri, USA). Promix 1, proteins standard mixture, was from LaserBio Labs (Sophia Antipolis, France). Bis[sulfosuccinimidyl] suberate (BS<sup>3</sup>) was from Thermo Scientific (Waltham Massachusetts, USA) and the K100 stabilization kit from CovalX (mix of three cross-linkers 1,1'-(suberoyldioxy)bisazabenzotriazole (SBAT), 1,1'-(suberoyldioxy)bisbenzotriazole (SBBT) and 1,1'-(glutaryldioxy)bisazabenzotriazole (GBAT)) (Schlieren, Switzerland). Trypsin Gold, Mass spectrometry Grade, was from Promega (Fitchburg, Wisconsin, USA). Dynabeads M280 streptavidin-coated magnetic beads were from Invitrogen (Carlsbad, New-Mexico, USA). The CPPs (R/W)<sub>9</sub>: Biot(O<sub>2</sub>)-G<sub>4</sub>-K-RRWWRRWRR-NH<sub>2</sub> (*m/z* 2127.12) and (R/W)<sub>16</sub>: Biot(O<sub>2</sub>)-G<sub>4</sub>-K-RRWWRRWRRWRRWRR-NH<sub>2</sub> (*m/z* 3309.76) were synthesized (Fmoc strategy) and purified in-house. Ziptip<sup>®</sup> C<sub>4</sub> pipette tips were from Millipore (Darmstadt, Germany).

### 2.2. Cross-linking reactions

Monomeric actin (G-actin) purchased lyophilized in Tris, ATP, and CaCl<sub>2</sub> was dialysed against a 20 mM HEPES, 150 mM NaCl, 0.5 mM DTT, 200 μM CaCl<sub>2</sub>, 200 μM ATP and 0.005% NaN<sub>3</sub> pH 8 buffer (Actin non-denaturing buffer). Stock solutions of CPP and actin were diluted in the actin non-denaturing buffer to obtain the desired concentrations (typically 10 μM). For CPP and albumin a 20 mM HEPES, 150 mM NaCl buffer was used. For all experiments, after 15 min pre-incubation of the CPP and protein partners, 50 mM cross-linkers solutions prepared *extemporaneously* in dimethylformamide (DMF) for the K100 kit, in water for BS<sup>3</sup>, were added to have a 2 mM final concentration of cross-linkers (200 folds higher than proteins concentration) and the reactions were allowed to proceed for 120 min at room temperature (RT) under gentle stirring. Typical cross-linking reactions (XL) volumes were 50 μL. For experiments with low peptides concentrations (1 or 0.1 μM), reaction volumes were adapted (200 μL and 2000 μL respectively) to allow purifying enough cross-linked peptides. Control reaction mixtures noted CT were performed in the same conditions but without cross-linkers. Cross-linking reactions were quenched by adding Tris-base (final concentration 15 mM).

### 2.3. Enzymatic trypsin digestion

25 μL of samples were submitted to in-solution tryptic digestion. Disulfide bridges were reduced (5 mM dithiothreitol) and cysteines alkylated (20 mM iodoacetamide). Trypsin digestion was conducted at 37 °C overnight (1:30 (w:w) protease-to-protein ratio).

### 2.4. Biotinylated cross-linked peptides affinity purification

A volume of tryptic digest corresponding to 80 pmol of CPP was incubated with 200 μg of streptavidin-coated magnetic beads for 60 min at RT under gentle stirring. Beads were conditioned before use and washed after incubation [12]. Finally, peptides were eluted from the beads with 5 μL 0.1 M hydrochloric acid (HCl) under gentle stirring (30 min).

### 2.5. MS analysis and data treatment

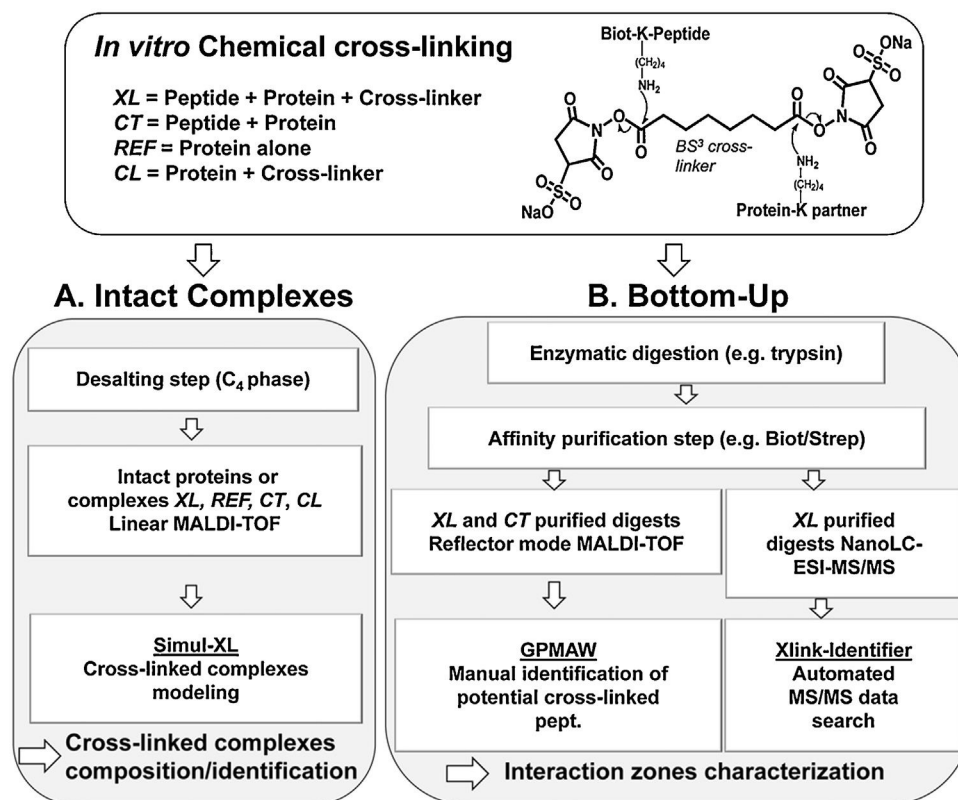
MALDI-TOF spectra were obtained with a MALDI-TOF/TOF AB4700 Proteomics Analyzer mass spectrometer (Applied Biosystems) in positive ions linear or reflector mode and delayed extraction. CHCA (Sigma) was used as the matrix and solubilized at 5 mg/mL in 1/1 ACN/0.1% TFA.

*Intact complexes analysis in linear mode MALDI-TOF*: 0.5 μL of non-digested reaction mixtures purified by C<sub>4</sub> Ziptip<sup>®</sup> were mixed with an equal volume of matrix and deposited on the sample holder. A total of 20,000 laser shots per sample were acquired in the *m/z* range 10,000–100,000 (focus mass 60,000).

Modelings of the spectra obtained for intact cross-linked complexes and corresponding controls were achieved using SIMUL-XL, a program developed in-house (Visual Basic V6.0).

*Bottom-up analysis in reflector mode MALDI-TOF/TOF*: After affinity purification, 0.5 μL of the elution mixture was mixed with 0.5 μL matrix and 0.5 μL were deposited on the sample holder. A total of 10,000 laser shots per spot were acquired in the *m/z* range 500–5000. External calibrations were realized using the peptide calibration standard Pepmix4 (Laser BioLabs). A manual interpretation of the spectra was done comparing the cross-linking reaction mixture (XL) and the peptide and protein mixture without cross-linker (CT) spectra. Differential peaks identified in XL spectrum were confronted with the cross-linked peptides lists generated *in silico* using General Protein Mass Analysis for Windows (GPMW) version 6.1 [19] (Lighthouse Data, Odense, Denmark) with a number of missed-cleavages for trypsin of 2.

*NanoLC-ESI-Orbitrap*: The complete characterization of the cross-linked peptides was performed using an Ultimate 3000 Nano-HPLC system (Dionex) coupled with a LTQ-Orbitrap-XL mass spectrometer (Thermo Scientific). Samples were injected by the autosampler and concentrated on a trapping column (Pepmap, C<sub>18</sub>, 300 μm × 5 mm, 5 μm, 100 Å, Dionex) with water containing 2% ACN and 0.1% formic acid (solvent A). After 10 min, the peptides were eluted onto the separation column (Pepmap, C<sub>18</sub>, 75 μm × 150 mm, 2 μm 100 Å, Dionex) equilibrated with 98% solvent A. Peptides were separated using the gradient 0–50 min 2–40% solvent B (98% ACN + 0.1% formic acid), 50–60 min 40–60% solvent B, and 60–70 min 60% solvent B at a flow rate of 200 nL/min. The LTQ-Orbitrap mass spectrometer is outfitted with a nano ESI interface. Electrospray emitters were 360/20 μm o.d. × 10 μm i.d. fused-silica tips (PicoTip Emitter, Standard Coated SilicaTip, New Objective). The heated capillary temperature and spray voltage were 200 °C and 1.5 kV, respectively. Orbitrap spectra (automated gain control (AGC) 2 × 10<sup>5</sup>) were collected from *m/z* 300–2000 at a resolution of 30,000 in the profile mode followed by data dependent sequential CID and HCD MS/MS spectra of the three



**Fig. 1 – General analytical workflow to in vitro identify and characterize protein partner(s) of a known CPP. The chemical cross-linking reaction between a biotinylated CPP, the BS<sup>3</sup> cross-linker and a potential protein partner is presented (top scheme). The cross-linked complexes were characterized using an intact complex analysis (A) and a bottom-up one (B).**

most intense ions with a normalized energy of 35 for both fragmentation modes. A dynamic exclusion time of 60 s was used to discriminate against previously analyzed ions.

MS/MS spectra were automatically searched using the Xlink-Identifier software [18]. Xlink-Identifier is a search engine for identifying and characterizing cross-linked peptides from label-free experiments (neither the peptide nor the cross-linker is isotopically labeled). It takes the MS/MS spectra in .dta or .mgf format and the protein sequences in FASTA format. Search parameters include the fragmentation technique, the nature of the dynamic modifications (carbamidomethylation (C), oxidation (M)), the maximum number of missed-cleavages, the precursor and fragment mass tolerance (respectively 10 ppm and 0.6 Da) and finally the definition of the cross-linker spacer arm mass involved in the formation of the cross-linking products (96.0211 Da for GBAT, 138.0618 Da for SBAT, SBBT or BS<sup>3</sup>).

As mentioned, Xlink-Identifier is a fully automated search engine specifically for cross-linking analysis. It is equipped with a visualization module allowing researchers to examine the annotated MS/MS spectra and details of each matched peak.

### 3. Results and discussion

In this study, we synthesized analogs of (R/W)<sub>9</sub> and (R/W)<sub>16</sub> CPPs (Biot(O<sub>2</sub>)-G<sub>4</sub>-K-RRWRRWRR-NH<sub>2</sub> and

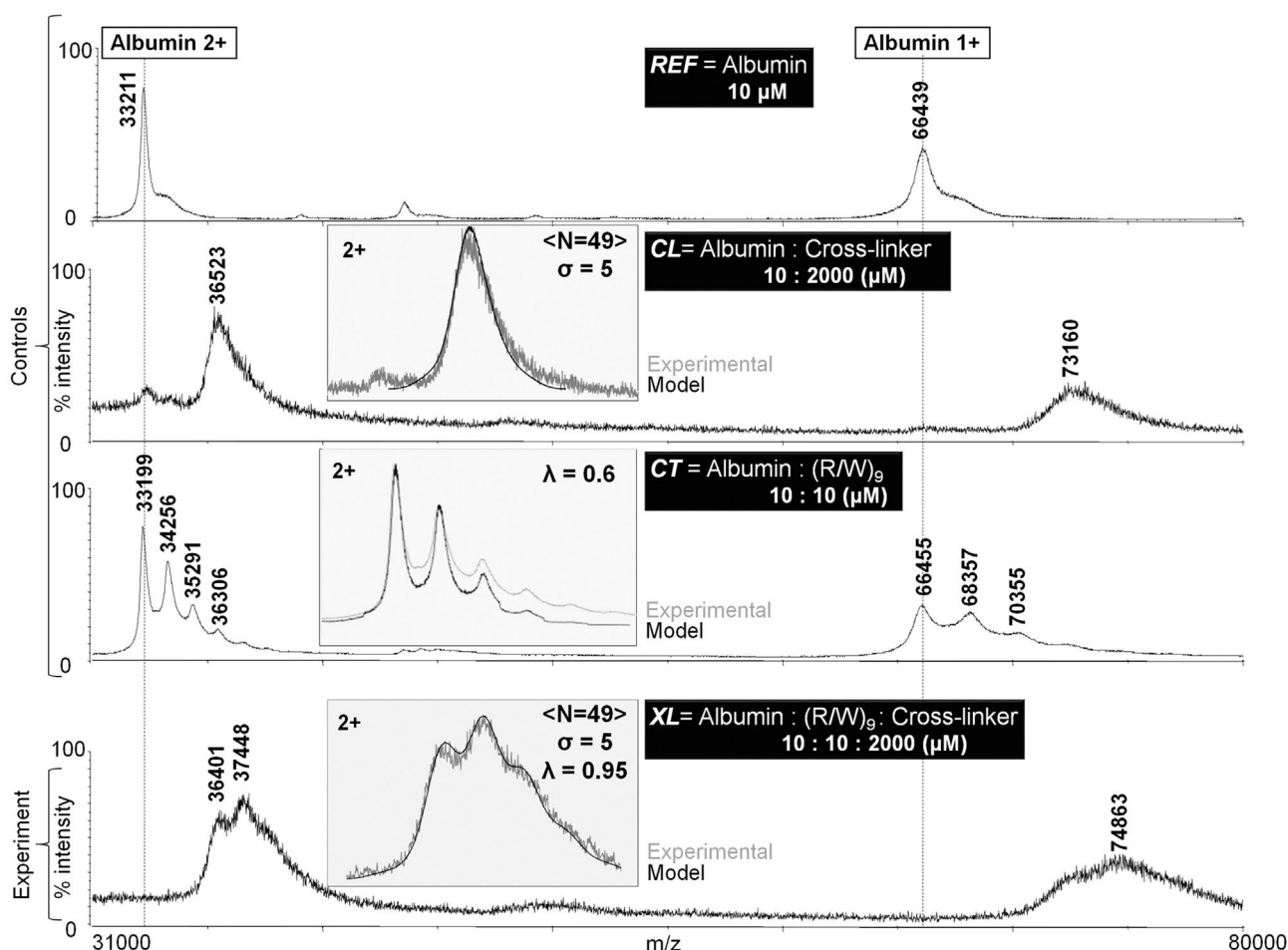
Biot(O<sub>2</sub>)-G<sub>4</sub>-K-RRWRRWRRWRRWRR-NH<sub>2</sub>, respectively) adapted for cross-linking experiments. To study the interaction of these CPPs with protein partners, a cross-linking-MS approach integrating two parallel strategies, an intact complex analysis and a bottom-up one, was developed and optimized (Fig. 1).

#### 3.1. Chemical cross-linking reaction

Actin is a globular 42 kDa protein that binds an ATP molecule and a divalent cation (Mg<sup>2+</sup> or Ca<sup>2+</sup>). These cofactors are necessary to maintain the integrity of the protein. Commercial actin was available as a lyophilized powder containing Tris buffer, which is not compatible with chemical cross-linking because of its reactive primary amine group. The sample was thus dialysed against a HEPES buffer also containing appropriate concentrations of salts and actin cofactors. Initial concentration of 10 μM was chosen for the two partners as the specificity of chemical cross-linking with activated esters was demonstrated in the low μM range [20].

We worked with a mixture of three cross-linkers (K100 stabilizing kit): SBAT, SBBT and GBAT (Fig. S-1). This mixture of cross-linkers is recommended to stabilize complexes with a molecular weight below 100 kDa and allows testing two spacer lengths at the same time (11.4 Å for SBAT and SBBT and 7.7 Å for GBAT). The K100 kit cross-linkers were solubilized in DMF. To make sure that the small proportion of this organic





**Fig. 2 – Positive ions linear mode MALDI-TOF of intact controls and chemical cross-linking reaction mixture.** The controls (REF, CL and CT) are presented on the three top spectra, the cross-linking experiment (XL) in the lower spectrum. Mass spectra were modeled using SIMUL-XL program. Modelings (black) that best fit the experimental data (gray) are presented in a box (doubly charged species).  $\sigma$ : standard deviation,  $\lambda$ : average number of  $(R/W)_9$ ,  $\langle N \rangle$ : average number of cross-linker.

solvent is not affecting the protein or peptide conformation and their interactions and thus the cross-linking reactions results, experiments were repeated with the BS<sup>3</sup> water soluble cross-linker which has the same spacer arm as SBAT or SBBT.

Supplementary Fig. 1 can be found, in the online version, at [doi:10.1016/j.euprot.2014.03.002](https://doi.org/10.1016/j.euprot.2014.03.002).

### 3.2. Intact complex analysis: determination of $(R/W)_9$ CPP – protein ratio in intact cross-linked complexes

**Linear mode MALDI-TOF:** MALDI-TOF is usually not a method of choice to study non-covalent complexes due to imperfect preservation of complexes during target preparation and unspecific non-covalent multimers desorption due to spatial proximity and high concentration in the matrix crystals. However it is well adapted to the study of covalent complexes formed using chemical cross-linking. Linear mode MALDI-TOF analyses of the intact complexes were carried out after a simple desalting step of the samples on a C<sub>4</sub> ZipTip®. The aim was to assess the efficiency of the cross-linking reaction

by determining the average number of  $(R/W)_9$  CPPs covalently attached per protein, for different reaction conditions.

To assess the contribution of the different components to the observed signal for a cross-linking reaction (XL), controls are necessary: the protein alone (REF), the protein plus the cross-linker only (CL) (to assess the mono-links, intra and inter cross-linked proteins), and the  $(R/W)_9$  CPP and the protein mixture without the cross-linker (CT) (to assess non-covalent interactions that could remain even with the acidic CHCA dissociative matrix). Many precautions need to be taken in the preparation of these controls and in mass spectra acquisition (same laser fluence and focus mass as for XL mixture). To quantitatively describe the different mass spectra obtained (number of mono-links, number of peptides covalently or non-covalently attached) a dedicated program SIMUL-XL was developed in our laboratory (Visual Basic V6.0) and is available as an open source on request at [gerard.bolbach@upmc.fr](mailto:gerard.bolbach@upmc.fr).

**Modeling of intact interacting species:** The modeling is based on the experimental peak profile of the protein alone (singly and multiply charged ions) (REF spectrum, Fig. 2). To each point of this profile a hypothetical distribution of  $(R/W)_9$  or mono-links or  $(R/W)_9$  + mono-links is added. The average number of

mono-links is noted  $\langle N \rangle$  and of  $(R/W)_9$  CPPs is  $\lambda$ . Mono-links for which the average number is much greater than one are added according to a gaussian law with a chosen standard deviation value  $\sigma$ , while  $(R/W)_9$  CPPs are added according to a poisson law since the average number  $\lambda$  is less than one.

The final peak profile is build up by superimposition of all these profiles (Fig. S-2). The comparison of this convoluted profile with the experimental data allows adjusting the hypothetical distribution of  $(R/W)_9$  or cross-linkers or  $(R/W)_9$  + cross-linkers. The initial profile of the protein alone is large including the isotopic pattern, initial axial velocity effects and matrix adducts. In comparison, the isotopic pattern of the  $(R/W)_9$  peptide and of the cross-linker are negligible and they are thus not taken into account. Simple gaussian and Poisson distributions were found to fit satisfactory the experimental data.

Supplementary Fig. 2 can be found, in the online version, at [doi:10.1016/j.euprot.2014.03.002](https://doi.org/10.1016/j.euprot.2014.03.002).

This work was performed for both interacting systems,  $(R/W)_9$ -albumin and  $(R/W)_9$ -actin, for two ratios 1:1 and 6:1. The 1:1 ratio is giving us an indication on the cross-linking reaction “yield” and the 1:6 ratio (large excess of peptide) was chosen to see if the CPP could have more than one interaction sites on the proteins. Spectra and modelings obtained for the  $(R/W)_9$ -albumin system in a 1:1 ratio are presented in Fig. 2.

The CL mixture was satisfactory fitted with a gaussian distribution of cross-linkers centered on  $\langle N = 49 \rangle$  with  $\sigma = 5$ . This suggests a very good accessibility of the lysine residues (total 52). The same distribution was found to fit all the charge states demonstrating that the covalent attachment is the major interaction and that non-covalent attachments are negligible [21] even in the presence of a large excess of cross-linkers.

The XL mixture spectrum was correctly fitted for all the charge states (+1 to +4) using the previous cross-linker gaussian distribution ( $\langle N = 49 \rangle$  and  $\sigma = 5$ ) and a Poisson distribution of  $(R/W)_9$  CPP with a mean value of  $\lambda = 0.95$ . Increasing the ratio  $(R/W)_9$ -albumin to 6:1, the mean value is  $\lambda = 3$  indicating the existence of several binding sites on the albumin protein (Fig. S-3).

Supplementary Fig. 3 can be found, in the online version, at [doi:10.1016/j.euprot.2014.03.002](https://doi.org/10.1016/j.euprot.2014.03.002).

For CT samples, we observed that peaks can be fitted with a CPP Poisson distribution with a mean value strongly dependent on the charge state: 0.7 for +1 and 0.5 for +3. Similar results were found with both CPPs-protein ratios 1:1 and 6:1. This observation is in good agreement with in-source dissociation of the non-covalent complexes in MALDI-TOF [21].

All these results clearly show that a high cross-linking yield is observed for the  $(R/W)_9$ -albumin interacting system and that artifacts due to non-covalent interactions, if any, are negligible for XL reaction mixture.

For the  $(R/W)_9$ -actin interacting system in a 1:1 ratio, CL was fitted with  $\langle N = 8 \rangle$  mono-links and  $\sigma = 2$ , indicating that some of the 17 lysines are likely less exposed than others to the cross-linkers. The best fit for the XL mixture was obtained with  $\langle N = 6 \rangle$  mono-links and  $\lambda = 0.3$  CPP (data not shown). Hence the cross-linking yield was low for the  $(R/W)_9$ -actin system. For a 6:1 ratio, the S/N was too low and mass spectra were thus not interpreted using SIMUL-XL program.

The results of this strategy devoted to intact complexes show an *in vitro* direct interaction of  $(R/W)_9$  CPP with both albumin and actin. It also easily gives access to the efficiency of the cross-linking reaction, by providing an estimation of the average number of accessible lysine residues and covalently attached CPPs. Cross-linked complexes obtained with a relatively high yield could correspond to a rather homogeneous or heterogeneous population. This was further studied using the bottom-up analysis.

### 3.3. Bottom-up analysis: characterization of the interacting zones between $(R/W)_9$ or $(R/W)_{16}$ CPPs and actin or albumin proteins

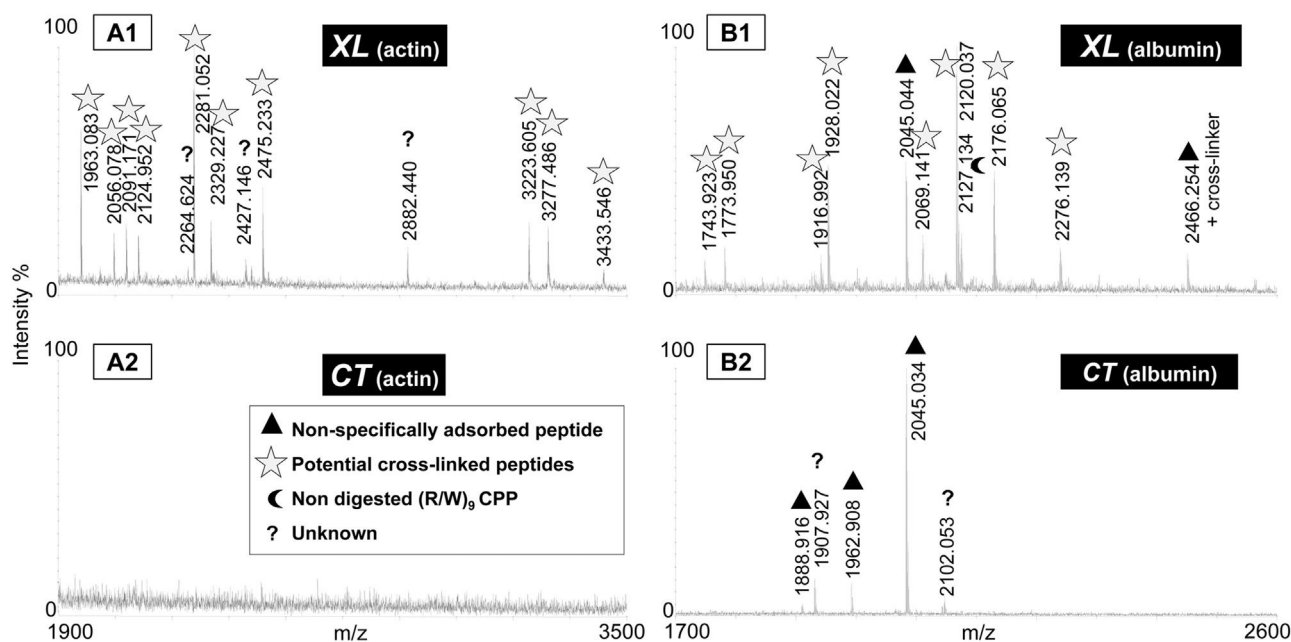
MALDI-TOF analysis of the purified cross-linking digests revealed a robust, reproducible and efficient method to evaluate and optimize the cross-linking experiments workflow (evaluation of the efficiency of the affinity purification step, the specificity of the cross-linking reactions, the influence of the CPP to protein ratio etc.).

*Purification of the biotinylated tryptic cross-linked peptides.* Among the possible products of the cross-linking reactions, we had to distinguish between the mono-linked species on one of the partner (CPP or protein) of the complex, the intra- or inter-protein cross-linked products and the cross-linked products involving both the CPP and the protein, the latter being those of interest. After in-solution trypsin digestion of the cross-linking (XL) and control (CT) mixtures, species containing the tryptic biotinylated CPP were selectively enriched through a biotin/streptavidin affinity purification step using streptavidin-coated magnetic beads. An improvement of the initial protocol [12] was achieved by adding a 10 min incubation step of the beads in 100% ACN to remove non-specifically adsorbed species. The elution step was also adapted to the cross-linking-MS workflow to allow the elution of the cross-linking products in an aqueous phase compatible with the nanoLC-MS/MS system. Three different acidic elution solutions were compared: 0.1% TFA, 10% formic acid (FA) and 0.1 M HCl, the latter turned to be the most efficient. MALDI-TOF analysis of the streptavidin beads directly spotted onto the sample holder after treatment with 0.1 M HCl revealed that the great majority or even the totality of the cross-linked species were released from the beads.

Fig. S-4 emphasizes the absolute necessity of a purification/enrichment step for the observation of subpicomole cross-linked products. Moreover this enrichment is highly selective as the MALDI-TOF analysis of the purified samples showed only few hydrophobic albumin tryptic peptides non-specifically adsorbed on the beads. It is important to notice that biotin/streptavidin affinity purification performed on intact complexes prior to trypsin digestion revealed unsuccessful. In this case, the recognition of the biotin tag by the streptavidin molecules is likely hindered when the CPPs (2.1 or 3.3 kDa) are associated to a bulky protein (42 kDa for actin and 66 kDa for albumin).

Supplementary Fig. 4 can be found, in the online version, at [doi:10.1016/j.euprot.2014.03.002](https://doi.org/10.1016/j.euprot.2014.03.002).

*Manual bottom-up MALDI-TOF data analysis.* For XL experiments of both CPP-actin and CPP-albumin interacting systems, a visual comparison of the MALDI-TOF spectra of



**Fig. 3 – Positive ions reflector mode MALDI-TOF of reaction mixtures after biotin/streptavidin affinity purification for actin-(R/W)<sub>9</sub> and albumin-(R/W)<sub>9</sub> interacting systems. XL stands for cross-linking experiments with actin or albumin (A1, B1), CT are the corresponding controls without cross-linking agents (A2, B2).**

XL and CT reactions, allowed to identify discriminating peaks appearing only in the XL reaction. Fig. 3 shows XL and CT MALDI spectra obtained for the (R/W)<sub>9</sub> CPP-actin or (R/W)<sub>9</sub> CPP-albumin systems. Same peaks were obtained with the (R/W)<sub>16</sub> CPP corresponding systems.

The manual confrontation of this list of discriminating peaks with the GPMW list of predicted cross-links led to the identification of cross-linked peptides (Table 1). Sequences of cross-linked peptides are presented in Table S-1.

Supplementary Table S-1 can be found, in the online version, at [doi:10.1016/j.euprot.2014.03.002](https://doi.org/10.1016/j.euprot.2014.03.002).

As external calibration of each spectrum was realized on the nearest position, the cross-linked peptides were identified with a very satisfying mass accuracy (<20 ppm) providing a good confidence in the results. For the (R/W)<sub>9</sub> or (R/W)<sub>16</sub>-actin interacting system a total of nine cross-linked peptides were reproducibly (five independent experiments) identified, which correspond to four different zones (six lysines) of the protein and for (R/W)<sub>9</sub> or (R/W)<sub>16</sub>-albumin, eight cross-linked peptides were identified as five different zones (seven lysines) of the protein.

**Specificity of the cross-linking reaction:** To look at the specificity of the reaction, we achieved a competition experiment with both actin and albumin. Therefore, 10  $\mu$ M of (R/W)<sub>9</sub> CPP were incubated with 10  $\mu$ M of each protein and 2 mM of BS<sup>3</sup> cross-linker. The MALDI-TOF spectra of the purified digests showed peaks corresponding to cross-linked peptides with actin and with albumin. Even if the intensity of the peaks could not be directly related to their abundance in MS, the close relative intensity of the peaks corresponding to cross-linked peptides with both proteins, strongly suggested that association constants of (R/W)<sub>9</sub> CPP for actin and albumin were within the same range of values.

The specificity of the cross-linking reaction was also checked by incubating (R/W)<sub>9</sub> CPP (10  $\mu$ M) with Promix1, a commercial mixture of three proteins: insulin (4  $\mu$ M), cytochrome C (12  $\mu$ M) and myoglobin (16  $\mu$ M) in the presence of the K100 kit of cross-linkers. The comparison of the MALDI-TOF spectra (reflector positive ions mode) of the cross-linking (XL) and control (CT) purified digests revealed no discriminating peaks. This is in good agreement with the absence of formation of complexes between (R/W)<sub>9</sub> CPP and these proteins, that was observed with the intact complex analysis (Fig. S-5). Actually, these three proteins insulin (pI=7.6), cytochrome C (pI=9.6) and myoglobin (pI=7.2) have either neutral or basic pI whereas actin (pI=5.2) and albumin (pI=5.8) both have acidic pI (UniprotKB database/ProtParam tool pI values). Therefore interactions between actin/albumin and the polycationic (R/W)<sub>9</sub> or (R/W)<sub>16</sub> CPPs are likely electrostatic, at least partly.

Supplementary Fig. 5 can be found, in the online version, at [doi:10.1016/j.euprot.2014.03.002](https://doi.org/10.1016/j.euprot.2014.03.002).

**MALDI-TOF/TOF characterization of the cross-linked peptides.** MALDI-TOF/TOF fragmentation of these ions gave mostly weak intensity and poor quality spectra. However in most spectra, identical series of peaks at *m/z* 259.1, 316.1, 373.1, 430.1 were observed corresponding to the N-terminal ions of the biotinylated tag added to the CPPs sequences (Biot(O<sub>2</sub>), Biot(O<sub>2</sub>)-G, Biot(O<sub>2</sub>)-GG and Biot(O<sub>2</sub>)-GGG respectively). This fragmentation highly competes with the formation of other fragments, which partly explains the poor quality of the high collision energy MS/MS spectra. However, the formation of these diagnostic ions constitutes a relevant marker of the presence of the biotinylated CPP in the cross-linked species.

The limited number of fragments obtained by MALDI-TOF/TOF did not allow characterization of the branched

**Table 1 – Manual interpretation of cross-linked peptides between (R/W)<sub>9</sub> and actin (A) or albumin (B) using GPMW software. Experimental versus simulated *m/z* values of identified cross-linked peptides between (R/W)<sub>9</sub> and actin or albumin obtained after trypsin digestion are presented in the two left columns. The delta mass is in the third column. Corresponding tryptic peptides of the protein and (R/W)<sub>9</sub> are in the two last columns. Sequences of the cross-linked peptides underlined were confirmed by MS/MS. Same results were obtained with (R/W)<sub>16</sub>.**

Exp. <i>m/z</i> of potential cross-linked peptides	Theo. <i>m/z</i> value (GPMW)	$\Delta$ mass (ppm)	Potential tryptic peptides involved in the complex	Potential tryptic (R/W) <sub>9</sub> involved in the complex
A. Actin-(R/W) <sub>9</sub> interacting system				
<u>1963.067</u>	<u>1963.079</u>	6.1	<u>322–332</u>	<u>1–7</u>
<u>2056.065</u>	<u>2056.086</u>	10.2	<u>202–210</u>	<u>1–7</u>
<u>2091.153</u>	<u>2091.174</u>	10.0	<u>322–333</u>	<u>1–7</u>
2124.952	2124.95	0.9	46–57	1–6
<u>2281.035</u>	<u>2281.051</u>	7.0	<u>46–57</u>	<u>1–7</u>
2329.208	2329.225	7.3	311–323	1–7
<u>2475.201</u>	<u>2475.24</u>	15.7	<u>308–321</u>	<u>1–7</u>
<u>3223.61</u>	<u>3223.653</u>	13.3	<u>91–111</u>	<u>1–7</u>
3277.543	3277.503	12.2	35–56	1–7
<u>3433.56</u>	<u>3433.605</u>	13.1	<u>35–57</u>	<u>1–7</u>
B. Albumin-(R/W) <sub>9</sub> interacting system				
<u>1743.916</u>	<u>1743.917</u>	0.5	<u>452–459</u>	<u>1–7</u>
<u>1773.943</u>	<u>1773.932</u>	6.2	<u>242–248</u>	<u>1–7</u>
<u>1916.992</u>	<u>1916.986</u>	3.1	<u>210–218</u>	<u>1–7</u>
<u>1928.016</u>	<u>1928.017</u>	0.5	<u>233–241</u>	<u>1–7</u>
2069.158	2069.142	7.7	548–557	1–7
2120.031	2120.03	0.5	24–34	1–6
or	<u>2120.03</u>	0.5	<u>25–34</u>	<u>1–7</u>
<u>2176.058</u>	<u>2176.049</u>	4.1	<u>35–44</u>	<u>1–7</u>
<u>2276.131</u>	<u>2276.131</u>	0.0	<u>24–34</u>	<u>1–7</u>
or	<u>2276.131</u>	0.0	<u>25–34</u>	<u>1–7</u>

peptides. Moreover, manual identification of cross-linked peptide can be envisaged for simple and well defined interacting system (e.g. two known partners) but is not suitable for more complex mixtures. Therefore we developed an online and automated separation (nanoLC system), data acquisition (ESI-LTQ-Orbitrap) and data search (Xlink-Identifier) adapted to more complex systems involving a known biotinylated peptide or protein with one or several unknown partners. This automated procedure was validated with the same CPPs-protein (actin or albumin) interacting systems.

**NanoLC-ESI-Orbitrap automatic data analysis.** NanoLC-ESI-MS/MS analyses of the purified digests were performed in order to confirm cross-linked peptides sequences. We noticed that these analyses of low abundant peptides had to be done rapidly after affinity purification to avoid sample adsorption on the plastic tubes. If the nanoLC-MS/MS analysis was done within the 2 days following the purification step then the sensitivity of the LTQ-Orbitrap mass spectrometer was adequate to detect cross-linked peptides and to trigger MS/MS on their precursor ions.

**MS/MS data treatment via Xlink-Identifier.** The MS/MS spectra generated were searched using Xlink-Identifier [18]. Xlink-Identifier generates a list of cross-linked peptides for a given mass tolerance. For our experiments the mass tolerance on the precursor ions was set at 10 ppm. For each cross-linked peptides identified, Xlink-Identifier indicates the position of the cross-linking sites and provides matching scores including XlinkScore and the mass accuracy for the precursor ions selected (Fig. S-6A and B). A link is also available to visualize the corresponding MS/MS spectrum with identified fragments

from both peptide chains (labeled a or b) and highlighted in different colors according to their type (b, y etc.).

Supplementary Fig. 6 can be found, in the online version, at [doi:10.1016/j.euprot.2014.03.002](https://doi.org/10.1016/j.euprot.2014.03.002).

NanoLC-ESI-MS/MS analysis of the cross-linked peptides allowed confirming the sequence of 7 out of the 10 cross-linked peptides previously identified for the CPPs-actin interacting systems in MALDI-TOF while all except one cross-linked peptides sequences were confirmed for the CPPs-albumin complex. Moreover three additional cross-linked peptides were found for the CPPs-albumin systems searching NanoLC-ESI-MS/MS data with Xlink-Identifier.

The identifications were mainly obtained using CID spectra and confirmed with HCD spectra (Fig. S-7). CID spectra were generally sufficient to validate the sequences.

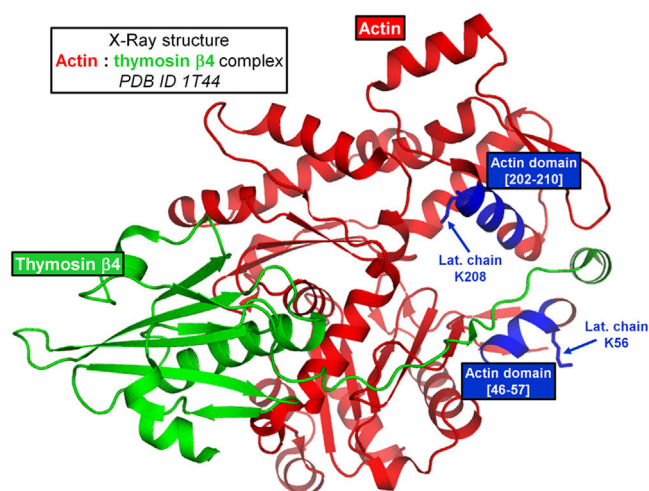
Supplementary Fig. 7 can be found, in the online version, at [doi:10.1016/j.euprot.2014.03.002](https://doi.org/10.1016/j.euprot.2014.03.002).

All the cross-linked peptides identified were the result of a covalent bond formation between the CPPs and actin or albumin via the longer version of the cross-linker (SBAT or SBBT). The similar results obtained using the BS<sup>3</sup> cross-linker showed that the leaving group type did not seem to modify the number or the nature of the cross-linked peptides obtained.

The vast majority of the cross-linked species identified between the biotinylated CPPs and the protein partner (albumin or actin) contained the biotinylated CPP under the Biot(O<sub>2</sub>)-G<sub>4</sub>-K-R form (residues 1–7) since the covalent link on the lysine prevents the trypsin cleavage after the K residue.

It is worth noticing that the covalent attachment of the Biot(O<sub>2</sub>)-G<sub>4</sub>-K-R group could be considered as a





**Fig. 4 – X-Ray structure of the complex actin: thymosin  $\beta$ 4 (PDB ID) [22].** The actin protein is colored in red, thymosin  $\beta$ 4 in green. Interaction zones characterized between the CPPs (R/W)<sub>9</sub> or (R/W)<sub>16</sub> and actin [202–210] from sub-domain 4 and [35–57] from sub-domain 2 are in blue (the region [40–50] is not visible on the crystal structure since it has not been crystallized). The lateral chain of the K residues K56 and K208 of actin, involved in the cross-linking are indicated.

post-translational modification of the peptides issued from the protein digestion and searched in this way with conventional search engine such as Mascot. However, contrary to Xlink-Identifier, with this type of search we can only have MS/MS information on the protein part and scoring of these cross-linked peptides is not properly made.

**Data rationalization in a biological context.** Cross-linking experiments were repeated several times for each interacting systems, giving reproducible results and leading to the identification of the same interaction zones for each system. For (R/W)<sub>9</sub> or (R/W)<sub>16</sub>-actin systems four different interaction zones were found: [35–57] containing the cross-linked K45 and K56, [91–111] containing the cross-linked K108, [202–210] containing the cross-linked K208 and [308–333] containing the cross-linked K310, K321 and K323 (Fig. S-8). For (R/W)<sub>9</sub> or (R/W)<sub>16</sub>-albumin interacting systems five different interaction zones were identified: [24–44], [210–218], [233–248], [452–459] and [548–557]. In addition, using the protein, actin or albumin, in a large excess (10:1 protein:CPP) gave the same results indicating that there would not be any preferential interaction site among the sites identified or any order in the colonization of these sites.

Supplementary Fig. 8 can be found, in the online version, at [doi:10.1016/j.euprot.2014.03.002](https://doi.org/10.1016/j.euprot.2014.03.002).

Using PyMOL free software program (DeLano Scientific LLC), it was possible to localize the identified interaction zones on the 3D structure (X-ray crystallography) of the protein for both systems. Demonstration of a direct interaction of both CPPs with the serum protein albumin is of interest for *in vivo* applications, since it might affect positively (protect from degradation) or negatively (sequestration of the CPP) the cell delivery of conjugated cargoes into cells [4]. This interaction

could involve electrostatic, hydrophobic or  $\pi$ /cation interactions, which could explain the dispersion and multiplicity of the interaction zones identified in albumin.

In the case of actin, three out of the four interaction zones containing the cross-linked lysines K45/K56, K108, K310/K321/K323, are located on the outer surface of the protein and are accessible to the CPPs (Fig. S-8). By contrast, the cross-linked lysine K208 is located in the close proximity of the nucleotide binding cleft and is less exposed to the solvent. It was previously reported that both (R/W)<sub>16</sub> and (R/W)<sub>9</sub> interact with G-actin [3]. We identified in this study identical domains of interaction for the two CPPs in actin. (R/W)<sub>16</sub> was shown previously to compete with the G-actin sequestering protein thymosin- $\beta$ 4 for binding to actin [3]. Interestingly, among the cross-linked regions identified in this study, the domains [35–57] (K45 and K56) and [202–210] (K208) are localized within the binding domain of thymosin- $\beta$ 4 [22] (Fig. 4), a spatial proximity that could explain the previous results of competition experiments. Actin polymerization/depolymerization dynamics is a complex process that involves numerous actin binding proteins (capping, nucleation, elongation, severing or bundle proteins). Kang and co-workers recently identified two discrete cation-binding sites within F-actin that they described as “polymerization” (driving actin filament assembly) and “stiffness” (modulating filament bending rigidity) sites [23]. Knowing that (R/W)<sub>9</sub> and (R/W)<sub>16</sub> CPPs can induce formation of stress fibers in cells that genetically lack F-actin formation, it is of particular interest to find herein that these CPPs bind G-actin within these two cation-binding regions that have been described to modulate actin dynamics in cells.

## 4. Conclusions

We set up and optimized a robust and efficient cross-linking-MS workflow allowing a complete characterization of *in vitro* CPP-protein interacting systems. This workflow originally designed for the study of CPP interaction partners can be applied to various interacting systems composed of two or more potential protein partners as soon as one of the two partners is known and can be labeled for enrichment purpose (e.g.: biotin tag). While the intact complex analysis gives reliable information on the cross-linking reaction yield and the cross-linked complexes composition according to the reaction conditions, the bottom-up analysis allows identification and characterization of the interaction zones.

Our study based on (R/W)<sub>9</sub> and (R/W)<sub>16</sub> CPPs illustrates the necessity of the use of cross-linking-MS for the *in vitro* study of interaction partners, and its complementarity with analytical techniques such as ITC and NMR.

Other CPPs can now be analyzed using the same strategy, in particular Tat that was shown to remodel actin cytoskeleton in actin-encapsulated giant vesicles [24]. Biologically speaking, we can anticipate that many, if not all, CPPs are not so inert and that similar cross-linking-MS strategies will be helpful to characterize membrane and intracellular partners for these peptides. Mapping the domains of interaction by cross-linking approaches will also help understanding, at least partly, the biological activity of CPPs.

## Conflict of interest

There is no conflict of interest.

## Transparency document

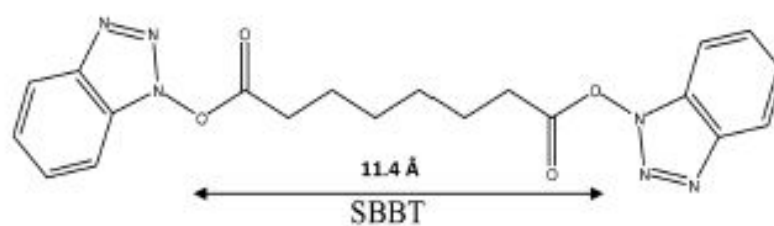
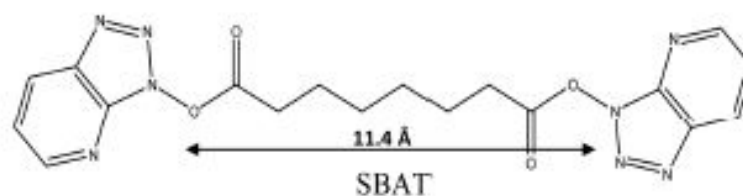
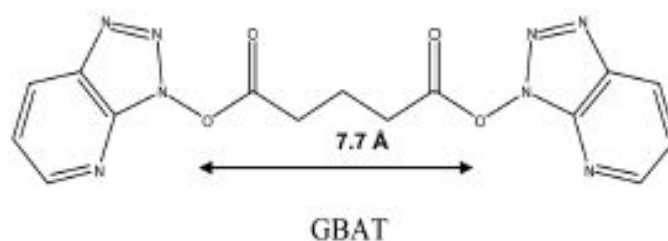
The [Transparency document](#) associated with this article can be found in the online version.

## REFERENCES

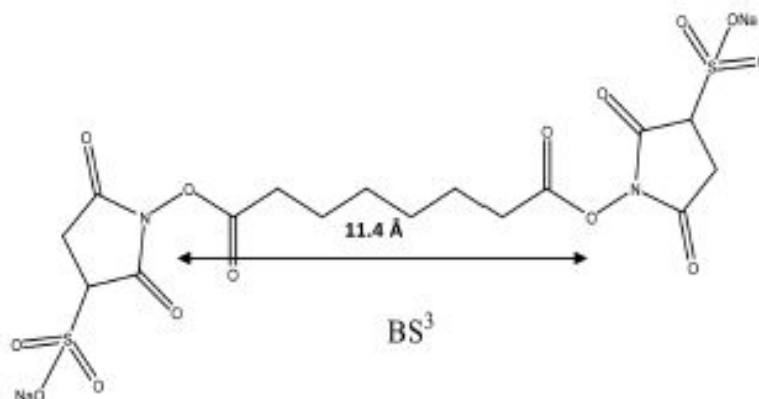
- [1] Bechara C, Sagan S. Cell-penetrating peptides: 20 years later, where do we stand? *FEBS Lett* 2013;587:1693–702.
- [2] Howl J, Matou-Nasri S, West DC, Farquhar M, Slaninová J, Ostenson C-G, et al. Bioportide: an emergent concept of bioactive cell-penetrating peptides. *Cell Mol Life Sci* 2012;69:2951–66.
- [3] Delaroche D, Cantrelle F-X, Subra F, Van Heijenoort C, Guittet E, Jiao C-Y, et al. Cell-penetrating peptides with intracellular actin-remodeling activity in malignant fibroblasts. *J Biol Chem* 2010;285:7712–21.
- [4] Kosuge M, Takeuchi T, Nakase I, Jones AT, Futaki S. Cellular internalization and distribution of arginine-rich peptides as a function of extracellular peptide concentration, serum, and plasma membrane associated proteoglycans. *Bioconjug Chem* 2008;19:656–64.
- [5] Mädler S, Bich C, Touboul D, Zenobi R. Chemical cross-linking with NHS esters: a systematic study on amino acid reactivities. *J Mass Spectrom* 2009;44:694–706.
- [6] Chowdhury SM, Du X, Tolić N, Wu S, Moore RJ, Mayer MU, et al. Identification of cross-linked peptides after click-based enrichment using sequential collision-induced dissociation and electron transfer dissociation tandem mass spectrometry. *Anal Chem* 2009;81:5524–32.
- [7] Sohn CH, Agnew HD, Lee JE, Sweredoski MJ, Graham RLJ, Smith GT, et al. Designer reagents for mass spectrometry-based proteomics: clickable cross-linkers for elucidation of protein structures and interactions. *Anal Chem* 2012;84:2662–9.
- [8] Tang X, Munske GR, Siems WF, Bruce JE. Mass spectrometry identifiable cross-linking strategy for studying protein–protein interactions. *Anal Chem* 2005;77:311–8.
- [9] Müller MQ, Zeiser JJ, Dreiocker F, Pich A, Schäfer M, Sinz A. A universal matrix-assisted laser desorption/ionization cleavable cross-linker for protein structure analysis. *Rapid Commun Mass Spectrom* 2011;25:155–61.
- [10] Ihling C, Schmidt A, Kalkhof S, Schulz DM, Stingl C, Mechtler K, et al. Isotope-labeled cross-linkers and Fourier transform ion cyclotron resonance mass spectrometry for structural analysis of a protein/peptide complex. *J Am Soc Mass Spectrom* 2006;17:1100–13.
- [11] Paramelle D, Cantel S, Enjalbal C, Amblard M, Forest E, Heymann M, et al. A new generation of cross-linkers for selective detection by MALDI MS. *Proteomics* 2009;9:5384–8.
- [12] Sachon E, Bolbach G, Chassaing G, Lavielle S, Sagan S. Cgamma H2 of Met174 side chain is the site of covalent attachment of a substance P analog photoactivable in position 5. *J Biol Chem* 2002;277:50409–14.
- [13] Rinner O, Seebacher J, Walzthoeni T, Mueller LN, Beck M, Schmidt A, et al. Identification of cross-linked peptides from large sequence databases. *Nat Methods* 2008;5:315–8.
- [14] Fritzschke R, Ihling CH, Götze M, Sinz A. Optimizing the enrichment of cross-linked products for mass spectrometric protein analysis. *Rapid Commun Mass Spectrom* 2012;26:653–8.
- [15] Leitner A, Walzthoeni T, Kahraman A, Herzog F, Rinner O, Beck M, et al. Probing native protein structures by chemical cross-linking, mass spectrometry, and bioinformatics. *Mol Cell Proteomics* 2010;9:1634–49.
- [16] Söderberg CAG, Lambert W, Kjellström S, Wiegandt A, Wulff RP, Månsson C, et al. Detection of crosslinks within and between proteins by LC–MALDI–TOF/TOF and the software FINDX to reduce the MSMS-data to acquire for validation. *PLoS ONE* 2012;7:e38927.
- [17] Rasmussen MI, Refsgaard JC, Peng L, Houen G, Højrup P. CrossWork: software-assisted identification of cross-linked peptides. *J Proteomics* 2011;74:1871–83.
- [18] Du X, Chowdhury SM, Manes NP, Wu S, Mayer MU, Adkins JN, et al. Xlink-identifier: an automated data analysis platform for confident identifications of chemically cross-linked peptides using tandem mass spectrometry. *J Proteome Res* 2011;10:923–31.
- [19] Nielsen T, Thaysen-Andersen M, Larsen N, Jørgensen FS, Houen G, Højrup P. Determination of protein conformation by isotopically labelled cross-linking and dedicated software: application to the chaperone, calreticulin. *Int J Mass Spectrom* 2007;268:217–26.
- [20] Nazabal A, Wenzel RJ, Zenobi R. Immunoassays with direct mass spectrometric detection. *Anal Chem* 2006;78:3562–70.
- [21] Sachon E, Clodic G, Blasco T, Bolbach G. Protein desolvation in UV matrix-assisted laser desorption/ionization (MALDI). *J Am Soc Mass Spectrom* 2007;18:1880–90.
- [22] Irobi E, Aguda AH, Larsson M, Guerin C, Yin HL, Burtinck LD, et al. Structural basis of actin sequestration by thymosin- $\beta$ 4: implications for WH2 proteins. *EMBO J* 2004;23:3599–608.
- [23] Kang H, Bradley MJ, McCullough BR, Pierre A, Grintsevich EE, Reisler E, et al. Identification of cation-binding sites on actin that drive polymerization and modulate bending stiffness. *Proc Natl Acad Sci USA* 2012;109:16923–7.
- [24] Mishra A, Lai GH, Schmidt NW, Sun VZ, Rodriguez AR, Tong R, et al. Translocation of HIV TAT peptide and analogues induced by multiplexed membrane and cytoskeletal interactions. *Proc Natl Acad Sci* 2011;108:16883–8.

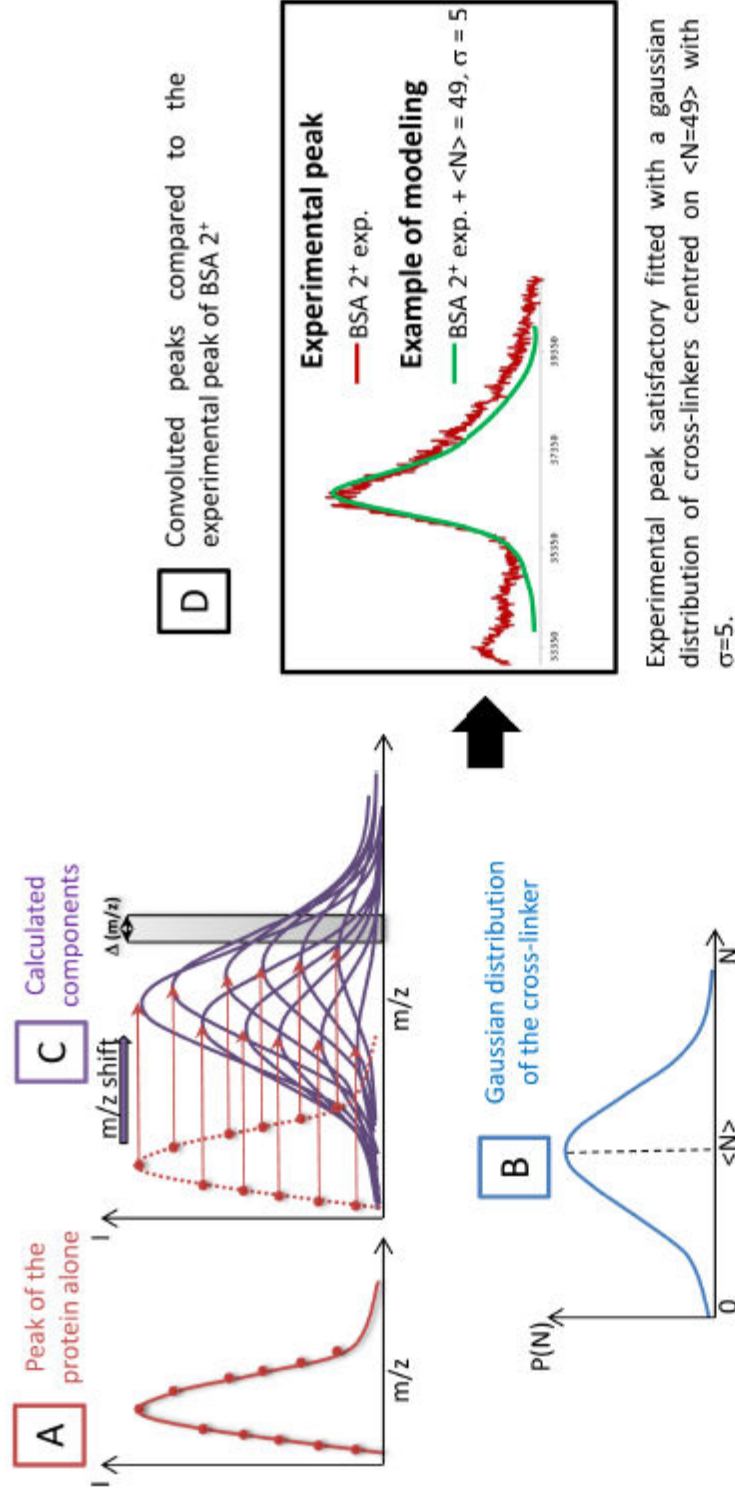
**Figure S-1:** Schematic view of the different cross-linkers used in the cross-linking experiments

**K100 stabilizer kit (CoValx)**



**Cross-linker BS<sup>3</sup> (Thermo Scientific)**



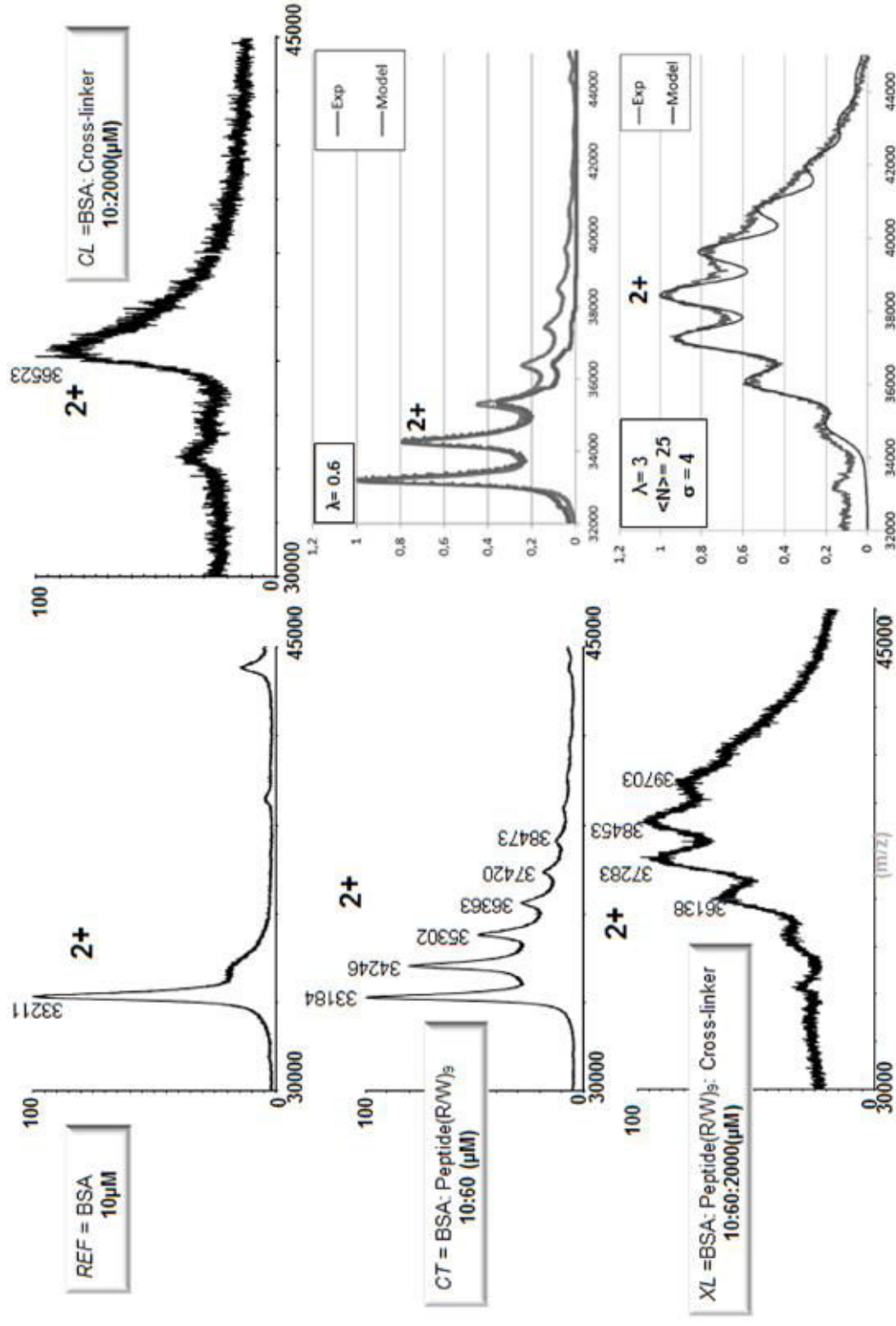


**Figure S-2. Modeling of the experimental peak.**

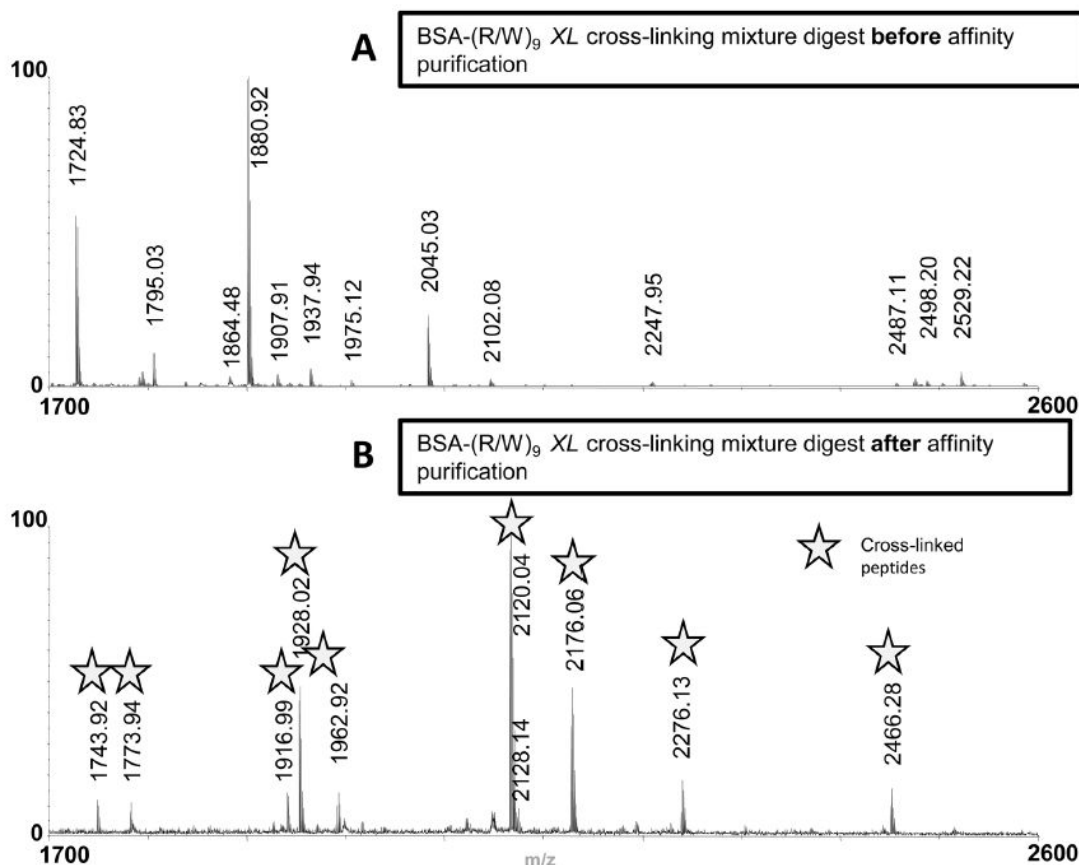
Modeling of the experimental peak of the protein+cross-linker or protein+peptide or protein+peptide+cross-linker. The starting point is the peak (with a charge  $z = 1$  or  $2 \dots$ ) of the protein alone (A). Depending on the peak width, a few tens to a few hundred points were considered. The hypothetical distribution of peptides or cross-linkers or peptides+cross-linkers is described by a gaussian or Poisson distribution. In the presented example (B), the distribution is a gaussian function,  $N$  is the number of cross-linkers and  $\langle N \rangle$  is the average number of cross-linkers. The different components can be calculated and are presented in (C). After summation of these components in a given window  $\Delta(m/z)$ , the convoluted peak can be compared to the experimental one. Then the distribution of peptides, cross-linkers or peptides+cross-linkers is adjusted and chosen to get the best fit (D). The width of the protein peak is large enough to ignore the effects of the isotopic pattern of the peptide and cross-linker.



**Figure S-3:** Linear MALDI-TOF spectra and modelings of the (R/W)<sub>9</sub>-albumin interacting system for a 6:1 pep: prot ratio



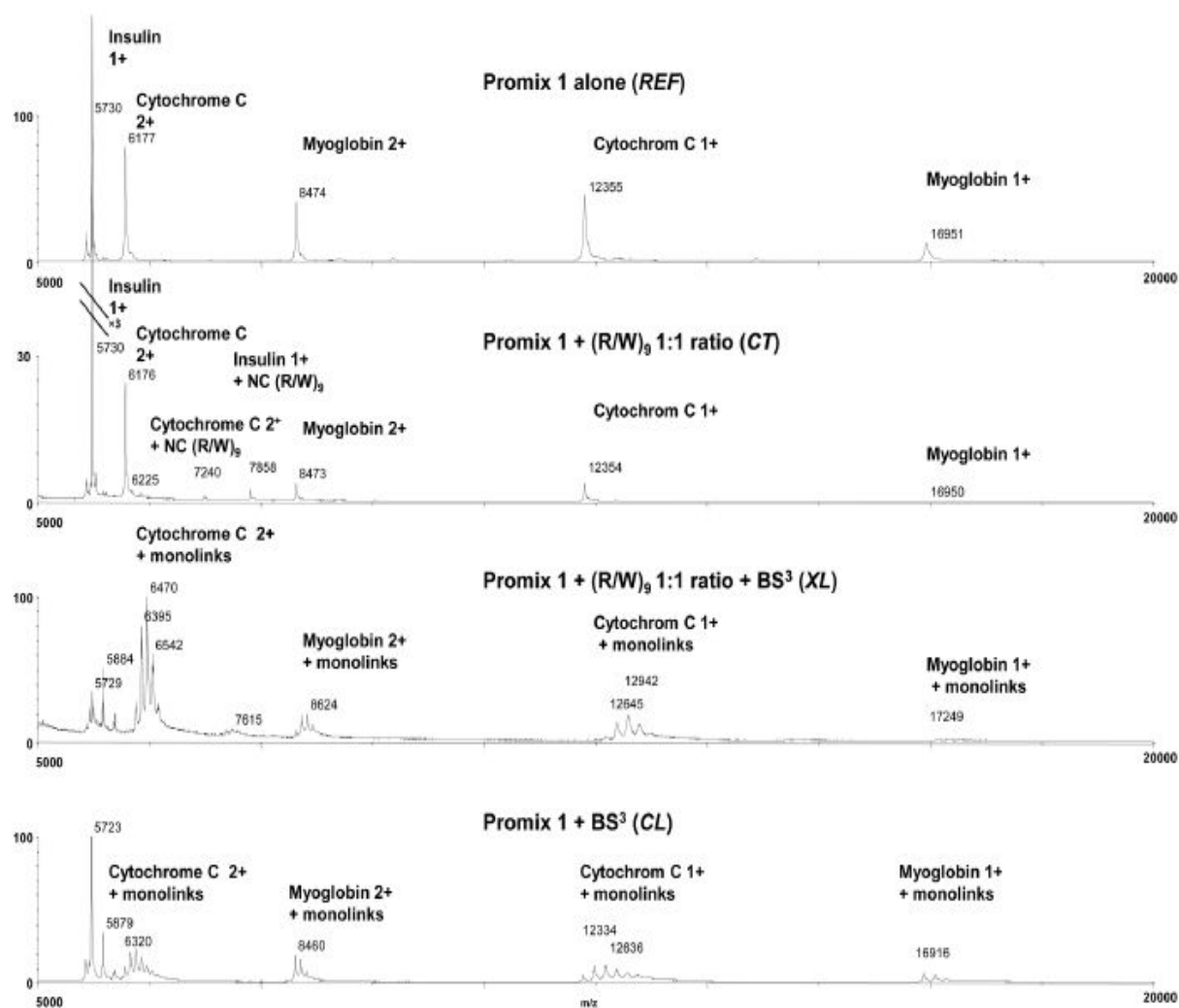
**Figure S-4:** Efficiency of the affinity purification step for the detection of cross-linked species: case of the BSA-(R/W)<sub>9</sub> XL cross-linking mixture.



Positive ions reflector mode MALDI-TOF of reaction mixtures before (A) and after (B) biot/strep affinity purification for the albumin-(R/W)<sub>9</sub> interacting system. XL stands for cross-linking experiment.

Stars represent the cross-linked peptides. No cross-linked peptides are observed without enrichment.

Figure S-5: Linear mode MALDI-TOF analysis of (R/W)<sub>9</sub>-Promix1 interacting system

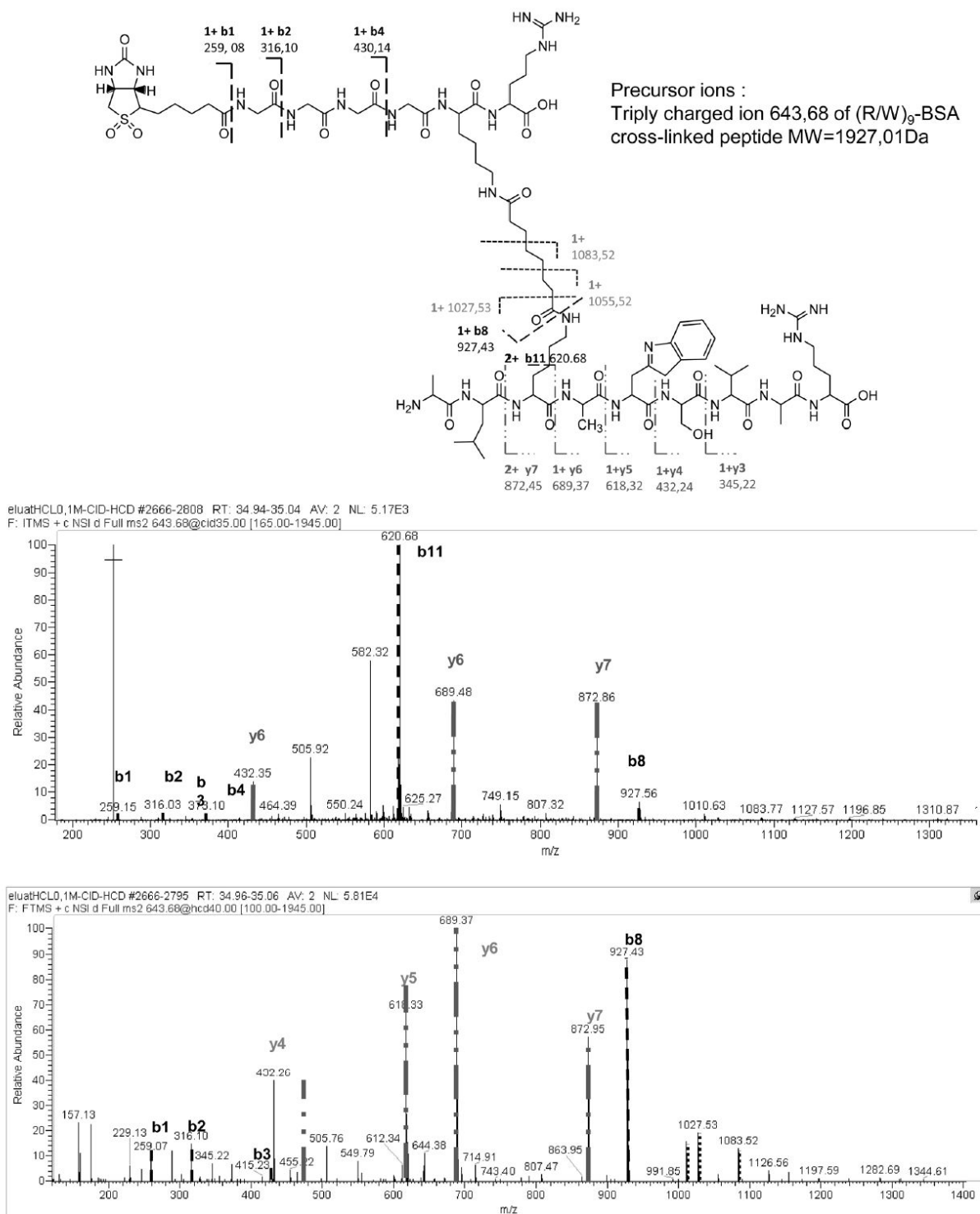


**Figure S-6: Xlink Identifier page results for the Biot(O<sub>2</sub>)G<sub>4</sub>-(R/W)<sub>9</sub>-albumin (A) or Biot(O<sub>2</sub>)G<sub>4</sub>-(R/W)<sub>9</sub>-Actin (B) interacting systems**

**A**

	Hit	CS	Score	Sp	Delta Mass (ppm)	Exp MS	Theo MS	E Ion	Y Ion	F Ion
2371	peptide [1079420] (ref) (IP: 551335.1) serum albumin precursor [Bos taurus] - Cell penetrating peptide [OTXK-SE2AEN - 805508*9]	5	39.542	156.197	0.046	2276.125	2276.125	17	20	/HMS: valantactio, IM-CID-STD, 2371, 3-2 E-63513355, 100011000, 38
2383	peptide [1079420] (ref) (IP: 551335.1) serum albumin precursor [Bos taurus] - Cell penetrating peptide [OTXK-SE2AEN - 805508*9]	5	37.788	145.586	1.254	2276.125	2276.125	14	19	/HMS: valantactio, IM-CID-STD, 2383, 1-1 E-25411557, 100011000, 38
2371	peptide [1079420] (ref) (IP: 551335.1) serum albumin precursor [Bos taurus] - Cell penetrating peptide [OTXK-SE2AEN - 805508*9]	5	37.656	157.174	0.066	2276.125	2276.125	15	21	/HMS: valantactio, IM-CID-STD, 2371, 5-2 E-13115402, 100011000, 38
2383	peptide [1079420] (ref) (IP: 551335.1) serum albumin precursor [Bos taurus] - Cell penetrating peptide [OTXK-SE2AEN - 805508*9]	5	36.208	145.450	1.254	2276.125	2276.125	14	18	/HMS: valantactio, IM-CID-STD, 2383, 1-4 E-82118581, 100011000, 38
2349	peptide [1079420] (ref) (IP: 551335.1) serum albumin precursor [Bos taurus] - Cell penetrating peptide [OTXK-SE2AEN - 805508*9]	5	21.140	47.571	-0.445	2120.024	2120.024	13	13	/HMS: valantactio, IM-CID-STD, 2349, 5-0 E-10191500, 100011000, 38
2357	peptide [1079420] (ref) (IP: 551335.1) serum albumin precursor [Bos taurus] - Cell penetrating peptide [OTXK-SE2AEN - 805508*9]	5	20.196	36.566	-0.524	2120.024	2120.024	10	15	/HMS: valantactio, IM-CID-STD, 2357, 1-1 E-11021005, 100011000, 38
2349	peptide [1079420] (ref) (IP: 551335.1) serum albumin precursor [Bos taurus] - Cell penetrating peptide [OTXK-SE2AEN - 805508*9]	4	44.443	142.700	-0.565	2276.127	2276.125	13	11	/HMS: valantactio, IM-CID-STD, 2349, 4-2 E-13314573, 100011000, 38
2357	peptide [1079420] (ref) (IP: 551335.1) serum albumin precursor [Bos taurus] - Cell penetrating peptide [OTXK-SE2AEN - 805508*9]	4	42.124	69.810	-0.537	2120.024	2120.024	14	13	/HMS: valantactio, IM-CID-STD, 2357, 4-0 E-7931650, 100011000, 38
2349	peptide [1079420] (ref) (IP: 551335.1) serum albumin precursor [Bos taurus] - Cell penetrating peptide [OTXK-SE2AEN - 805508*9]	4	38.468	121.760	-0.588	2276.127	2276.125	12	14	/HMS: valantactio, IM-CID-STD, 2349, 4-1 E-30491572, 100011000, 38
2357	peptide [1079420] (ref) (IP: 551335.1) serum albumin precursor [Bos taurus] - Cell penetrating peptide [OTXK-SE2AEN - 805508*9]	4	36.825	54.885	-1.250	2120.027	2120.024	11	9	/HMS: valantactio, IM-CID-STD, 2357, 4-3 E-27912587, 100011000, 38
2349	peptide [1079420] (ref) (IP: 551335.1) serum albumin precursor [Bos taurus] - Cell penetrating peptide [OTXK-SE2AEN - 805508*9]	4	36.024	62.121	-0.587	2120.028	2120.024	12	14	/HMS: valantactio, IM-CID-STD, 2349, 4-0 E-13115402, 100011000, 38
2350	peptide [1079420] (ref) (IP: 551335.1) serum albumin precursor [Bos taurus] - Cell penetrating peptide [OTXK-SE2AEN - 805508*9]	4	32.248	59.578	-1.280	2120.027	2120.024	10	11	/HMS: valantactio, IM-CID-STD, 2350, 4-2 E-84115130, 100011000, 38
2483	peptide [1079420] (ref) (IP: 551335.1) serum albumin precursor [Bos taurus] - Cell penetrating peptide [OTXK-SE2AEN - 805508*9]	4	20.094	39.766	-1.466	2176.047	2176.043	8	12	/HMS: valantactio, IM-CID-STD, 2483, 4-4 E-29496527, 100011000, 38
2489	peptide [1079420] (ref) (IP: 551335.1) serum albumin precursor [Bos taurus] - Cell penetrating peptide [OTXK-SE2AEN - 805508*9]	4	24.021	8.115	-0.455	2176.044	2176.043	10	13	/HMS: valantactio, IM-CID-STD, 2489, 4-10 E-23446133, 100011000, 38
2484	peptide [1079420] (ref) (IP: 551335.1) serum albumin precursor [Bos taurus] - Cell penetrating peptide [OTXK-SE2AEN - 805508*9]	3	62.851	140.024	-0.091	2176.048	2176.043	14	13	/HMS: valantactio, IM-CID-STD, 2484, 3-3 E-21451168, 100011000, 38
2740	peptide [1079420] (ref) (IP: 551335.1) serum albumin precursor [Bos taurus] - Cell penetrating peptide [ALP-MSTVA - 805508*9]	3	58.484	144.646	-1.110	1828.013	1828.011	11	13	/HMS: valantactio, IM-CID-STD, 2740, 3-7 E-13314573, 100011000, 38
2489	peptide [1079420] (ref) (IP: 551335.1) serum albumin precursor [Bos taurus] - Cell penetrating peptide [OTXK-SE2AEN - 805508*9]	3	47.256	122.413	-1.468	2176.048	2176.043	10	12	/HMS: valantactio, IM-CID-STD, 2489, 3-4 E-30491572, 100011000, 38
2733	peptide [1079420] (ref) (IP: 551335.1) serum albumin precursor [Bos taurus] - Cell penetrating peptide [ALP-MSTVA - 805508*9]	3	37.492	87.423	-1.872	1828.015	1828.011	8	9	/HMS: valantactio, IM-CID-STD, 2733, 3-4 E-44811315, 100011000, 38
2481	peptide [1079420] (ref) (IP: 551335.1) serum albumin precursor [Bos taurus] - Cell penetrating peptide [1526*238 - 805508*9]	3	36.760	123.429	-1.731	1773.929	1773.926	9	9	/HMS: valantactio, IM-CID-STD, 2481, 3-9 E-27912587, 100011000, 38
2420	peptide [1079420] (ref) (IP: 551335.1) serum albumin precursor [Bos taurus] - Cell penetrating peptide [1526*238 - 805508*9]	3	36.488	133.400	-0.862	1773.927	1773.926	8	8	/HMS: valantactio, IM-CID-STD, 2420, 3-1 E-10316131, 100011000, 38
2540	peptide [1079420] (ref) (IP: 551335.1) serum albumin precursor [Bos taurus] - Cell penetrating peptide [1526*238 - 805508*9]	3	32.102	89.913	-1.740	1742.914	1742.911	9	12	/HMS: valantactio, IM-CID-STD, 2540, 3-4 E-13115402, 100011000, 38
2476	peptide [1079420] (ref) (IP: 551335.1) serum albumin precursor [Bos taurus] - Cell penetrating peptide [1526*238 - 805508*9]	3	30.265	44.111	-0.243	2122.015	2122.011	11	9	/HMS: valantactio, IM-CID-STD, 2476, 3-2 E-20941544, 100011000, 38
2572	peptide [1079420] (ref) (IP: 551335.1) serum albumin precursor [Bos taurus] - Cell penetrating peptide [1526*238 - 805508*9]	3	29.930	99.205	-1.424	1916.903	1916.900	10	5	/HMS: valantactio, IM-CID-STD, 2572, 3-5 E-33446132, 100011000, 38
1545	peptide [1079420] (ref) (IP: 551335.1) serum albumin precursor [Bos taurus] - Cell penetrating peptide [1526*238 - 805508*9]	3	27.946	86.591	-1.425	1742.914	1742.911	11	13	/HMS: valantactio, IM-CID-STD, 1545, 3-4 E-13314573, 100011000, 38
2384	peptide [1079420] (ref) (IP: 551335.1) serum albumin precursor [Bos taurus] - Cell penetrating peptide [OTXK-SE2AEN - 805508*9]	3	26.027	24.626	-1.042	2120.027	2120.024	8	8	/HMS: valantactio, IM-CID-STD, 2384, 3-0 E-15418516, 100011000, 38
2480	peptide [1079420] (ref) (IP: 551335.1) serum albumin precursor [Bos taurus] - Cell penetrating peptide [1526*238 - 805508*9]	3	25.870	56.875	-4.123	2092.923	2092.914	6	9	/HMS: valantactio, IM-CID-STD, 2480, 3-2 E-20449773, 100011000, 38
1564	peptide [1079420] (ref) (IP: 551335.1) serum albumin precursor [Bos taurus] - Cell penetrating peptide [1526*238 - 805508*9]	3	25.240	85.200	-1.835	1916.903	1916.900	8	7	/HMS: valantactio, IM-CID-STD, 1564, 3-2 E-84115130, 100011000, 38
2349	peptide [1079420] (ref) (IP: 551335.1) serum albumin precursor [Bos taurus] - Cell penetrating peptide [OTXK-SE2AEN - 805508*9]	3	21.552	21.688	-0.947	2120.028	2120.024	7	8	/HMS: valantactio, IM-CID-STD, 2349, 3-1 E-33446132, 100011000, 38
2422	peptide [1079420] (ref) (IP: 551335.1) serum albumin precursor [Bos taurus] - Cell penetrating peptide [1526*238 - 805508*9]	3	19.316	52.411	-0.243	2122.015	2122.011	8	12	/HMS: valantactio, IM-CID-STD, 2422, 3-5 E-15418516, 100011000, 38

**Figure S-7:** Example of the MS/MS characterization of a cross-linked peptide

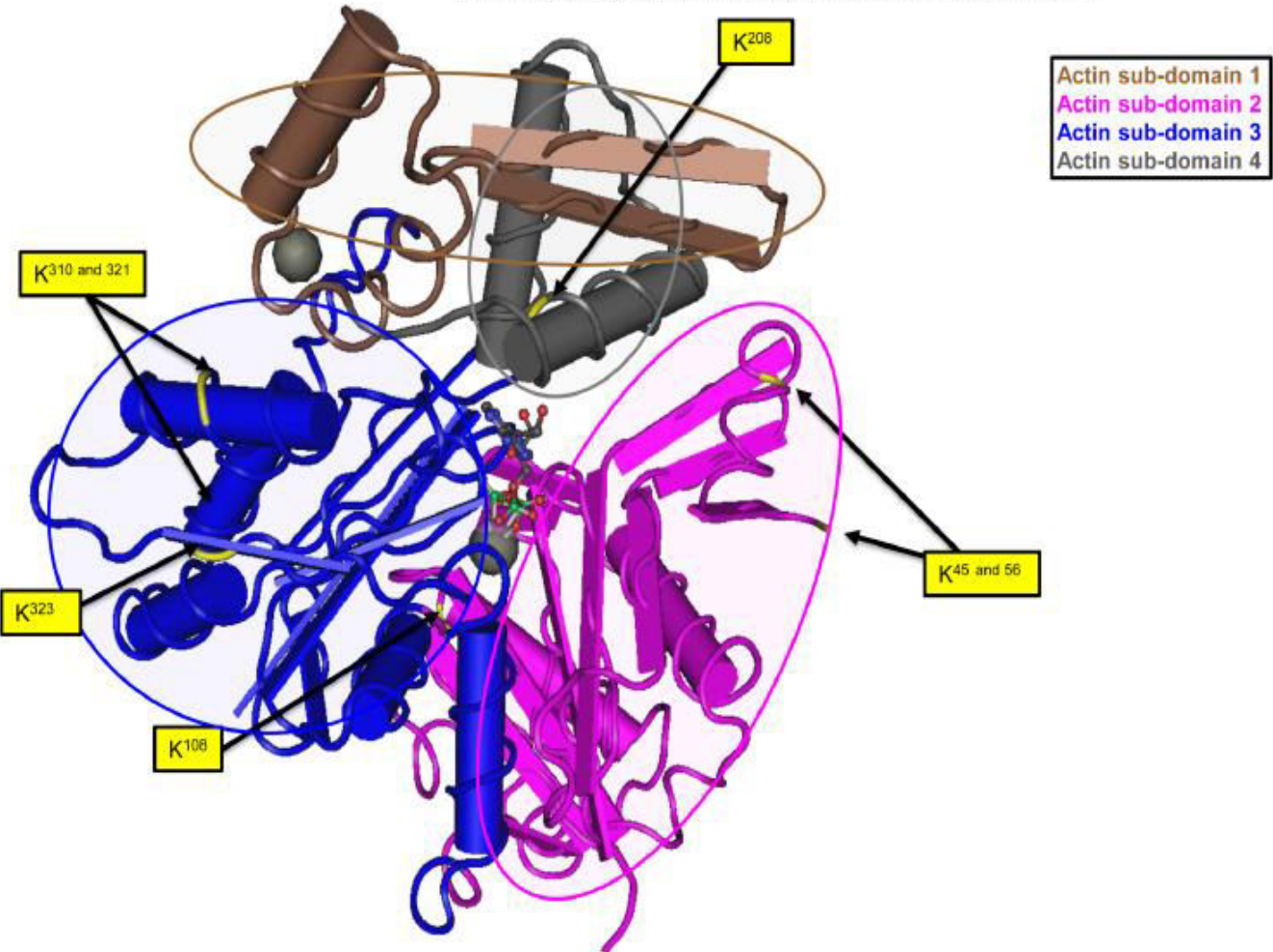


**Figure S-8: Actin protein (X-Ray crystallography, PDB ID 3EL2)**

Chemically cross-linked lysines are presented in yellow.

The 4 actin sub-domains are circled.

K<sup>310</sup> and K<sup>321</sup> from the observed actin tryptic sequence <sup>308</sup>MQKEITALAPSTMK<sup>321</sup> Actin sub-domain 3  
K<sup>323</sup> from the observed actin tryptic sequence <sup>322</sup>IKLIAPPER<sup>353</sup> Actin sub-domain 3  
K<sup>108</sup> from the observed actin tryptic sequence <sup>91</sup>VAPEEHPTLLTEAPLNPKANR<sup>111</sup> Actin sub-domain 2  
K<sup>45</sup> and K<sup>56</sup> from the observed actin tryptic sequence <sup>35</sup>HQGVVMVGMGQKDSYVGDEAQS<sup>57</sup> Actin sub-domain 2  
K<sup>208</sup> from the observed actin tryptic sequence <sup>202</sup>EIVRDIKEK<sup>210</sup> Actin sub-domain 4



## Chapitre 5 : Vers la capture *in cellulo* de partenaires d'interaction du peptide vecteur (R/W)<sub>9</sub> par photocross-linking et spectrométrie de masse

L'idée à l'origine de cette approche est la suivante : « Puisque le peptide (R/W)<sub>9</sub> est un peptide vecteur (ou plus communément CPP pour *cell penetrating peptide*), profitons de sa capacité à s'internaliser dans les cellules pour réaliser la capture de ses partenaires dans les conditions « réelles » c'est-à-dire « *in cellulo* » .

Les expériences préalables réalisées en pontage covalent ou cross-linking chimique pour caractériser l'interaction des peptides (R/W)<sub>9</sub> et (R/W)<sub>16</sub> avec l'actine et l'albumine *in vitro* (partenaires potentiels au vue de la littérature [2]) ont mis en évidence l'intérêt de cette technique pour la détermination des partenaires et des zones d'interaction (chapitre 4). Cependant, elles nous ont également fait prendre conscience d'un certain nombre de difficultés à surmonter pour appliquer ce type de stratégie à des mélanges protéiques complexes : faible rendement des réactions de cross-linking (nécessité d'avoir une cinétique de la réaction de pontage en accord avec la durée de vie des complexes), diversité des espèces cross-linkées formées (diversité des protéines partenaires ciblées mais aussi hétérogénéité des sites de pontage sur une même protéine, d'où un signal dilué et globalement faible) et hydrolyse rapide des groupements réactifs des cross-linkers bifonctionnels.

De plus, pour être en mesure de capturer des interactions *in cellulo*, il faut bien entendu que les cross-linkers puissent franchir la barrière plasmique des cellules, ce qui n'est pas le cas pour beaucoup d'entre eux, et qu'ils arrivent intacts dans le milieu intracellulaire, c'est-à-dire sans avoir subi d'hydrolyse de leurs fonctions réactives.

Ainsi, il est rapidement apparu que le cross-linking chimique ne pourrait pas être utilisé pour capturer des partenaires d'interactions *in cellulo*. C'est pour cette raison que nous nous sommes tournés vers le photocross-linking qui permet de pallier le problème d'hydrolyse des groupements réactifs et de contrôler précisément le moment où la réaction a lieu, simplement en réalisant une irradiation UV de l'échantillon.

### I. Mise au point et validation biologique d'une version photoactivable du CPP (R/W)<sub>9</sub> : photo(R/W)<sub>9</sub>

#### 1. Design et synthèse de photo(R/W)<sub>9</sub>

Ne nous intéressant qu'aux partenaires du peptide (R/W)<sub>9</sub> et non à l'ensemble des interactions pouvant se faire dans la cellule nous avons choisi de placer un groupement photoactivable de type benzophénone directement sur le peptide vecteur. Il a été incorporé en N-ter de la séquence (R/W)<sub>9</sub> *via* l'addition d'un résidu lysine dont la fonction amine primaire de la chaîne latérale a été fonctionnalisée avec un groupement benzophénone [277], [278]. De manière à pouvoir extraire le



## Résultats : Vers la capture de partenaires par photocross-linking in cellulo

peptide (R/W)<sub>9</sub> ainsi que les partenaires d'interaction auquel il se serait lié de manière covalente, nous avons choisi de placer une étiquette de type biotine à l'extrémité N-terminale de ce peptide vecteur photoactivable. Afin que la biotine ne vienne pas interférer avec l'activité biologique de la séquence (R/W)<sub>9</sub>, un bras espaceur composé de cinq glycines a été introduit entre la biotine et la lysine porteuse de la sonde photoactivable.

Ainsi la structure choisie pour le peptide photoactivable appelé photo(R/W)<sub>9</sub> est présentée Figure 32.

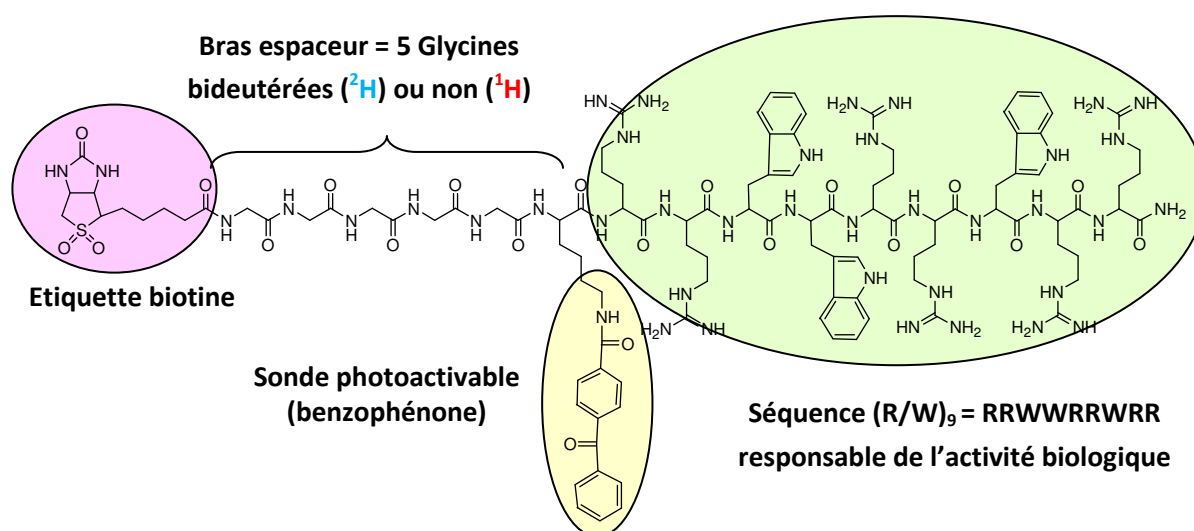


Figure 32: Structure développée du peptide photo(R/W)<sub>9</sub> mettant en évidence les fonctions de ses différentes parties constitutives.

Le bras espaceur constitué de 5 glycines bideutérées ou non est utile pour la quantification du peptide internalisé comme nous le verrons dans la partie I.2.b mais permet également de repérer facilement les espèces photocross-liquées lors de l'analyse par spectrométrie de masse [279]. En effet, en utilisant un mélange équimolaire des formes deutérées et non deutérées du peptide photo(R/W)<sub>9</sub>, les espèces contenant le peptide photo(R/W)<sub>9</sub> apparaîtront sur les spectres de masse sous forme de doublets d'égales intensités séparés de 10 Da (et donc normalement sans recouvrement de leurs massifs isotopiques) comme illustré Figure 33.

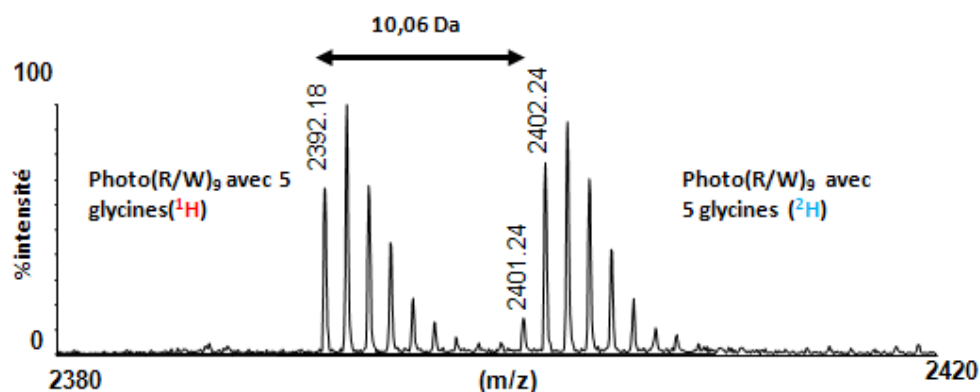


Figure 33 : Signal obtenu en spectrométrie de masse MALDI-TOF pour un mélange 1/1 des versions deutérées et non deutérées du peptide photo(R/W)<sub>9</sub>.



La synthèse et la purification de ces peptides ont été réalisées par Rodrigue Marquant, membre de l'équipe 1 du Laboratoire des BioMolécules (UMR 7203).

Une fois le peptide synthétisé et avant de commencer à réaliser des expériences de photocross-linking *in cellulo*, il nous a fallu évaluer l'effet de la présence de la sonde photoactivable sur la cytotoxicité du peptide photo(R/W)<sub>9</sub> ainsi que sur sa capacité d'internalisation dans les cellules, en comparaison avec le CPP (R/W)<sub>9</sub>. Nous avons également dû vérifier que le peptide photo(R/W)<sub>9</sub> conservait la capacité de remodelage de l'actine observée précédemment pour le CPP (R/W)<sub>9</sub>.

## 2. Validation biologique de photo(R/W)<sub>9</sub>

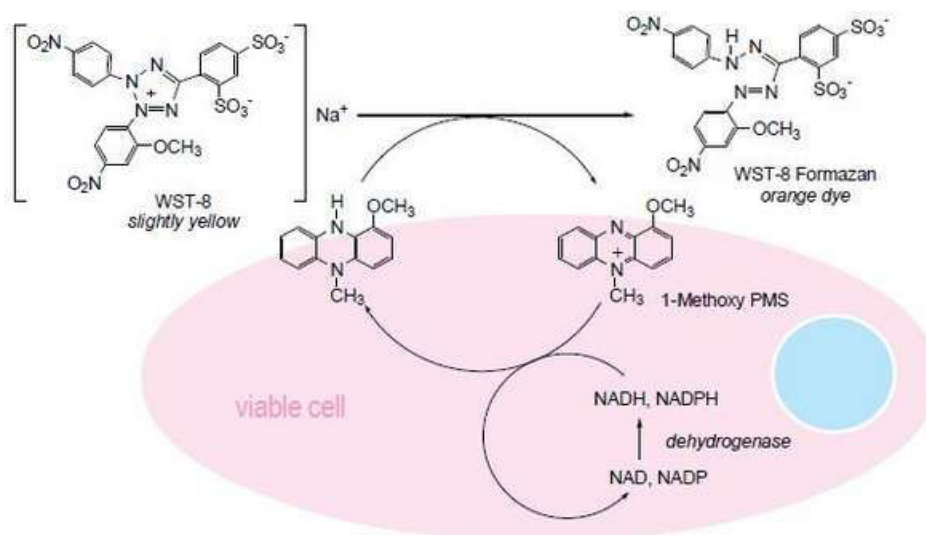
Pour tester l'effet du peptide photo(R/W)<sub>9</sub>, nous avons à chaque fois comparé les résultats obtenus à ceux du peptide (R/W)<sub>9</sub> de référence (thèse de Diane Delaroche). La séquence de ce **peptide (R/W)<sub>9</sub> de référence** noté (R/W)<sub>9</sub> réf est Biot-G<sub>4</sub>-(R/W)<sub>9</sub> (Biot-G<sub>4</sub>-RRWWRRWRR).

### a. Evaluation de la cytotoxicité du peptide vecteur photo(R/W)<sub>9</sub>

L'évaluation de la cytotoxicité du peptide photo(R/W)<sub>9</sub> a été réalisée avec un test colorimétrique appelé CCK-8 (Cell Counting Kit-8) commercialisé par Dojindo.

#### Principe de la technique

Le kit CCK-8 est basé sur l'utilisation du composé WST-8 de couleur légèrement jaune qui est réduit en un composé coloré orangé, le formazan, en présence de cellules viables selon le mécanisme décrit ci-dessous (Figure 34). La formation de ce composé orangé et l'absorbance du milieu sont proportionnelles à l'activité des déshydrogénases, elle-même proportionnelle au nombre de cellules vivantes.



**Figure 34 : Mécanisme de transformation du WST-8 (composé légèrement jaune) en WST-8 formazan (composé orange) en présence de cellules viables** (Figure d'après le site dojindo.com).

Pour évaluer la cytotoxicité d'un peptide, celui-ci est incubé à différentes concentrations avec une quantité donnée de cellules (typiquement, dans notre cas 40 000 cellules). La mesure d'absorbance  $A$  pour les différentes concentrations de peptide est réalisée à une longueur d'onde de 450 nm. La valeur obtenue est comparée à un contrôle négatif  $A(ctl\ neg)$  correspondant à la solution de peptide

## Résultats : Vers la capture de partenaires par photocross-linking in cellulo

seul à la plus forte concentration testée et à un contrôle positif A(ctrl pos) correspondant aux cellules en l'absence de peptide.

Le calcul du pourcentage de viabilité cellulaire pour une concentration donnée est le suivant:

$$\% \text{viabilité} = \left( 1 - \frac{A(\text{ctrl pos}) - A(\text{conc testée})}{A(\text{ctrl pos}) - A(\text{ctrl nég})} \right) \times 100$$

A(ctrl pos) = Absorbance mesurée dans le puits contenant uniquement des cellules.

A(ctrl nég) = Absorbance mesurée dans le puits pour le peptide à la plus forte concentration testée en l'absence de cellules.

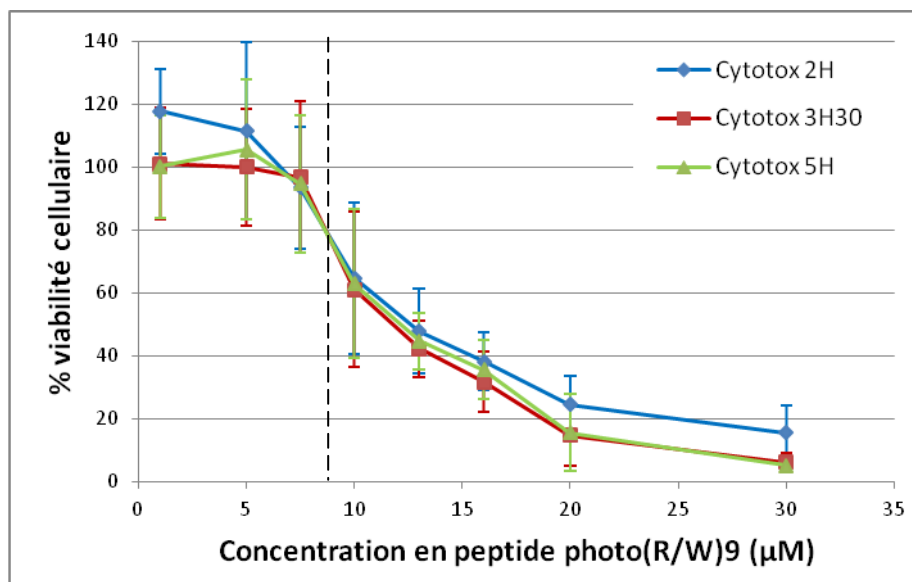
A(conc testée) = Absorbance mesurée dans le puits pour les cellules incubées avec une concentration testée de peptide.

### Expériences réalisées

La cytotoxicité du peptide photo(R/W)<sub>9</sub> pour les cellules EF a été mesurée après 2 h, 3h30 et 5 h d'incubation avec des concentrations en peptides comprises entre 1 et 30 µM (cf. partie matériel et méthodes III.1). La cytotoxicité du peptide photo(R/W)<sub>9</sub> vis-à-vis des cellules non-tumorales 3T3 a également été évaluée et comparée avec celle du peptide (R/W)<sub>9</sub>réf. Des triplicats ont été réalisés pour chaque condition d'une expérience et chaque expérience a été reproduite au minimum deux fois.

### Résultats et commentaires

Le résultat du test de cytotoxicité du peptide photo(R/W)<sub>9</sub> sur les cellules EF est présenté ci-dessous (Figure 35). Il indique que **le peptide photo(R/W)<sub>9</sub> est cytotoxique lorsqu'il est utilisé à une concentration supérieure à 7,5 µM.**



**Figure 35 : Pourcentage de viabilité des cellules EF à l'issue de 2 h, 3h30 et 5 h d'incubation avec le peptide photo(R/W)<sub>9</sub> pour des concentrations comprises entre 1 et 30 µM.** NB : La présence de points correspondant à une viabilité cellulaire supérieure à 100 % pour une concentration en peptide inférieure à 7,5 µM est liée essentiellement à l'erreur expérimentale sur le nombre de cellulesensemencées dans les puits.

La valeur seuil de cytotoxicité de photo(R/W)<sub>9</sub> pour les cellules 3T3 est identique à celle obtenue pour les cellules EF (pas montré).

Pour le peptide (R/W)<sub>9</sub> réf aucune cytotoxicité n'a été observée jusqu'à 30 µM.

Ainsi il apparaît que la sonde photoactivable de type benzophénone augmente la cytotoxicité du peptide puisqu'on passe d'une absence de cytotoxicité jusqu'à 30 µM pour le peptide (R/W)<sub>9</sub> réf à un seuil de cytotoxicité de 7,5 µM pour le peptide photo(R/W)<sub>9</sub>.

Ce seuil de cytotoxicité est supérieur à la concentration pour laquelle le remodelage de l'actine a été observé avec le peptide (R/W)<sub>9</sub> réf. Ainsi il nous faut nous assurer que le peptide photo(R/W)<sub>9</sub> s'internalise au moins aussi bien que le peptide (R/W)<sub>9</sub> réf et qu'il conserve une action sur le cytosquelette d'actine à une concentration inférieure à son seuil de cytotoxicité.

### **b. Evaluation de la capacité d'internalisation du peptide vecteur photo(R/W)<sub>9</sub>**

Pour évaluer la capacité d'internalisation du peptide photo(R/W)<sub>9</sub> dans les cellules EF nous avons fait appel à une méthode de quantification par spectrométrie de masse MALDI-TOF mise au point au laboratoire [256].

#### **Principe de la méthode**

Cette technique de quantification du peptide internalisé dans des cellules repose sur l'utilisation d'un étalon interne marqué isotopiquement par incorporation de glycines bideutérées au lieu des glycines non deutérées au niveau du bras espaceur. Ce marquage isotopique confère à ce peptide étalon des propriétés physico-chimiques identiques à celles du peptide non deutéré correspondant. Ces deux peptides ont ainsi des comportements identiques lors de l'étape de purification par affinité et ils vont être ionisés et détectés de la même façon en spectrométrie de masse.

Le schéma ci-dessous (Figure 36) reprend les différentes étapes du protocole (*cf.* partie matériel et méthodes III.2). Tout d'abord, le peptide est incubé à une concentration donnée (ex : 10 µM) avec un nombre de cellules connu (ex : 10<sup>6</sup> cellules). Cette incubation est généralement réalisée à la température de 37 °C (température physiologique) pour laquelle les différents mécanismes d'internalisation dépendants ou non d'un apport d'énergie sont possibles. Ce protocole a également été optimisé à 4 °C pour permettre d'évaluer uniquement la contribution des mécanismes d'internalisation indépendants de l'énergie (translocation directe). En effet, à 4 °C, la cellule ne peut plus synthétiser d'ATP.

A la fin de l'incubation, les cellules sont lavées et le peptide lié à la membrane (non internalisé) est digéré à la trypsine. Dans le cas d'une incubation à 4 °C, la trypsine est remplacée par de la pronase, enzyme qui conserve son activité à basse température. Une quantité connue d'étalon deutéré est alors introduite dans un volume connu de tampon de lyse. Il est important d'introduire l'étalon le plus tôt possible dans la séquence d'événements afin qu'il subisse les mêmes étapes que le peptide non marqué, en particulier cela permet de suivre les éventuelles dégradations du peptide. Les lysats cellulaires obtenus sont incubés avec des billes magnétiques greffées de streptavidine afin de récupérer par affinité les peptides internalisés et les peptides étalons qui sont tous deux porteurs d'une étiquette biotine. Ces billes sont par la suite incubées en conditions acides avec de la matrice CHCA ce qui permet l'élution des peptides biotinylés et le dépôt direct des éluats sur une plaque MALDI pour leur analyse en spectrométrie de masse.

## Résultats : Vers la capture de partenaires par photocross-linking in cellulo

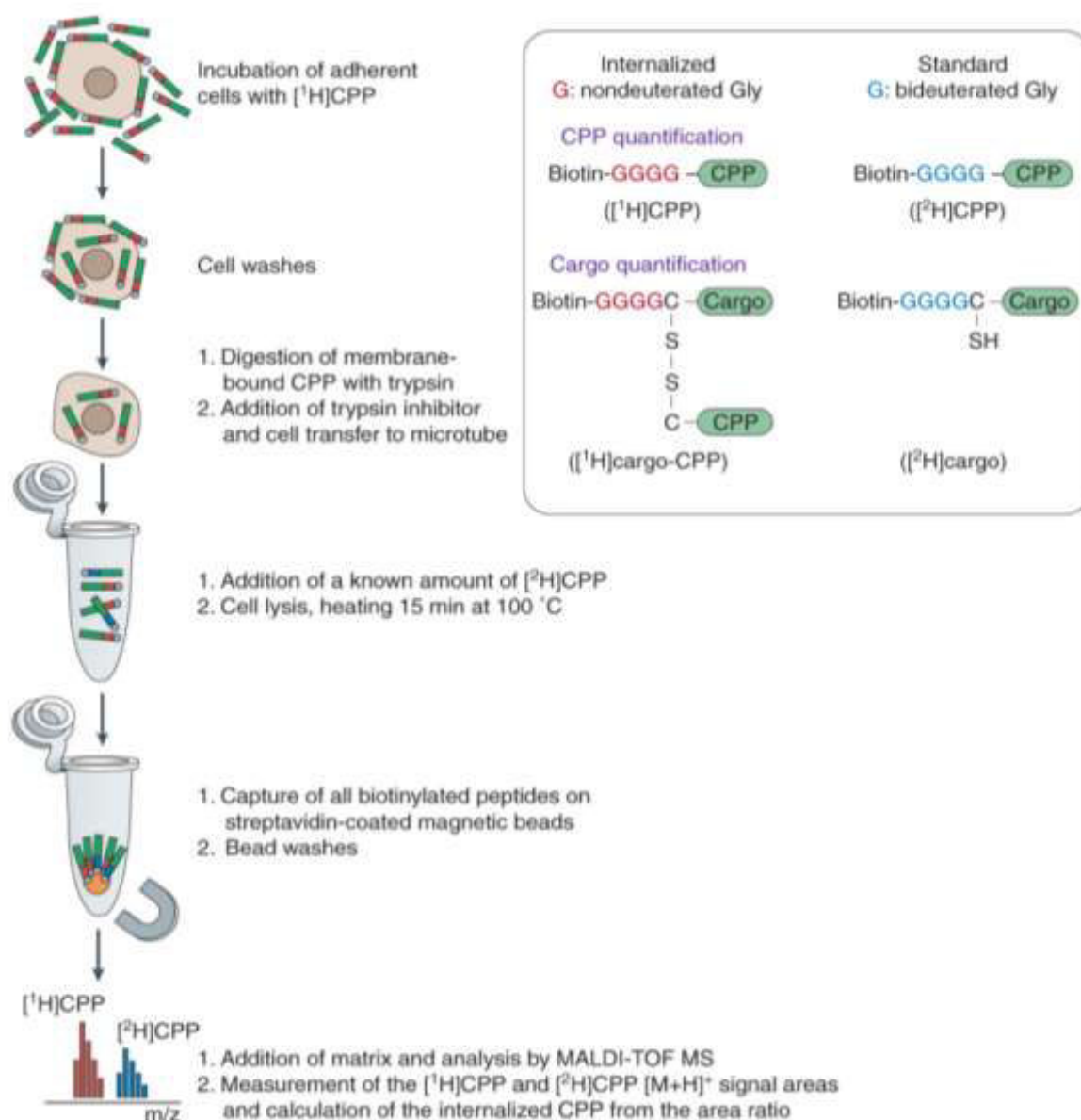


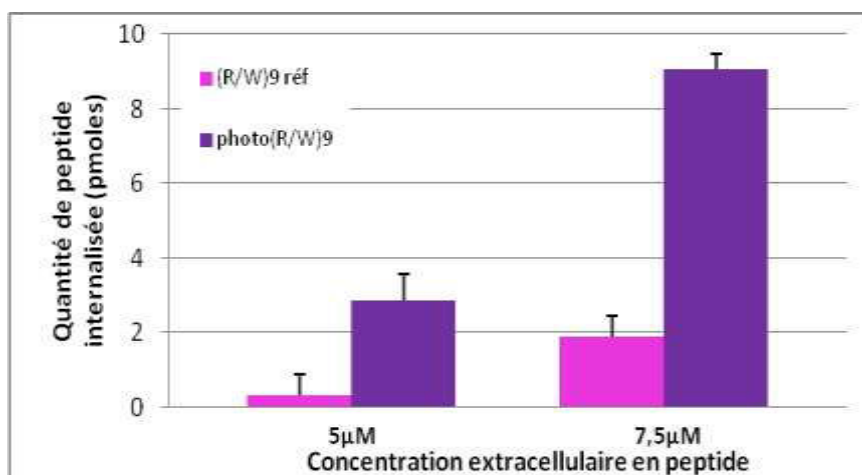
Figure 36 : Les différentes étapes du protocole de quantification de l'internalisation d'un peptide vecteur (CPP) (Figure tirée de [256]).

### Expériences réalisées

La capacité d'internalisation du peptide photo(R/W)<sub>9</sub> dans les cellules EF a été évaluée à 37 °C à l'issue d'une incubation de 75 min avec des concentrations extracellulaires en peptides égales à 5 et 7,5 µM. La quantification du peptide (R/W)<sub>9</sub> réf internalisé a été réalisée en parallèle pour ces mêmes concentrations en guise de contrôle.

### Résultats et commentaires

Le graphique ci-après (Figure 37) récapitule les résultats obtenus pour les expériences d'internalisation des peptides photo(R/W)<sub>9</sub> et (R/W)<sub>9</sub> réf réalisées sur les cellules EF.



**Figure 37 : Résultats obtenus pour les expériences de quantification de l'internalisation des peptides photo(R/W)<sub>9</sub> et (R/W)<sub>9</sub> réf dans les cellules EF à 37 °C pour des concentrations extracellulaires en peptides de 5 et 7,5 μM et une période d'incubation de 75 min.**

Tout d'abord, ces résultats montrent que le peptide photo(R/W)<sub>9</sub> a une capacité d'internalisation dans les cellules EF supérieure d'un facteur 5 à 10 par rapport à celle du peptide (R/W)<sub>9</sub> réf pour des concentrations extracellulaires de 7,5 et 5 μM, respectivement.

Ainsi, **la cytotoxicité augmentée du peptide photo(R/W)<sub>9</sub> par rapport au peptide (R/W)<sub>9</sub> réf pourrait être compensée dans nos expériences, par une meilleure capacité d'internalisation.** L'hydrophobie importante de la sonde benzophénone confère au peptide une meilleure affinité pour la membrane plasmique et peut expliquer, au moins partiellement, sa meilleure capacité d'internalisation.

Ensuite, ces résultats indiquent l'influence importante de la concentration extracellulaire en peptides sur la quantité de peptides internalisée. En effet, le peptide photo(R/W)<sub>9</sub> s'internalise environ 3 fois mieux à 7,5 μM (≈ 9 pmol internalisées) qu'à 5 μM (≈ 3 pmol internalisées). Cependant, étant donné le seuil de cytotoxicité du peptide photo(R/W)<sub>9</sub>, nous nous placerons par la suite à une concentration extracellulaire en peptides égale à 5 μM permettant d'avoir après 75 min environ 3 pmol de peptide photo(R/W)<sub>9</sub> internalisé dans 10<sup>6</sup> cellules soit une concentration intracellulaire d'environ 2 μM (volume de 10<sup>6</sup> cellules de l'ordre de 1,5 μL).

Par ailleurs, les expériences d'internalisation des peptides photo(R/W)<sub>9</sub> et (R/W)<sub>9</sub> réf réalisées avec les cellules 3T3 ont donné des résultats similaires indiquant que ces peptides s'internalisent aussi bien dans les cellules non tumorales 3T3 que dans les cellules tumorales EF.

### **c. Vérification de l'effet de remodelage de l'actine du peptide photo(R/W)<sub>9</sub> par immunofluorescence**

Nous avons vu précédemment que le peptide photo(R/W)<sub>9</sub> présentait une cytotoxicité supérieure à celle du peptide (R/W)<sub>9</sub> réf mais que cette toxicité plus importante pouvait être compensée par une capacité d'internalisation nettement supérieure, ce qui permet théoriquement, en jouant sur les concentrations extracellulaires de peptides, d'avoir une quantité équivalente de peptide internalisé. Il nous faut à présent vérifier que le peptide photo(R/W)<sub>9</sub> présente une activité biologique sur les cellules tumorales EF comparable à celle observée préalablement pour le peptide (R/W)<sub>9</sub> réf ainsi que

## Résultats : Vers la capture de partenaires par photocross-linking in cellulo

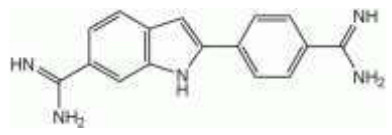
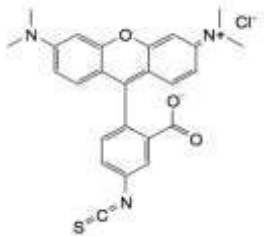
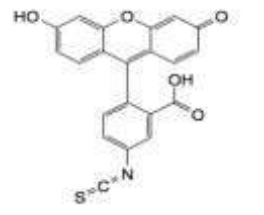
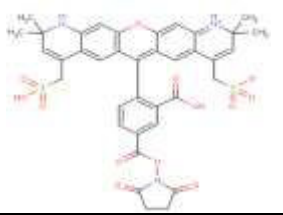
pour le peptide (R/W)<sub>16</sub> à savoir sa capacité à remodeler le cytosquelette d'actine, caractérisée par un retour des fibres de stress dans les cellules.

### Principe de l'immunofluorescence

L'immunofluorescence est une technique utilisée pour visualiser au sein de coupes de tissus ou de cellules fixées des molécules cibles telles que des protéines, des glycanes ou des petites molécules d'origine biologique ou non. Ces molécules cibles sont reconnues par des anticorps spécifiques qui sont associés à des fluorophores (groupements capables d'émettre de la lumière après excitation). Après marquage, les échantillons sont observés grâce à un microscope à fluorescence. La technique d'immunofluorescence peut être primaire ou secondaire. Le terme d'immunofluorescence fait par extension également référence au marquage et à la détection de molécules cibles par des ligands de haute affinité intrinsèquement fluorescents (ex : DAPI (4',6-diamidino-2-phénylindole) pour le marquage de l'ADN) ou pouvant être greffés par un fluorophore (ex : phalloïdine qui lie fortement l'actine F).

Parmi les fluorophores les plus communément utilisés en immunofluorescence se trouvent la FITC, le TRITC ou encore l'Alexa Fluor 568 (plus résistant au *photobleaching* c'est-à-dire à la destruction photochimique du fluorophore) dont les structures et les caractéristiques spectrales sont présentées (Tableau 6).

**Tableau 6 : Exemples de fluorophores communément utilisés en immunofluorescence.**

Nom du fluorophore	Structure	$\lambda_{\text{excitation}}$ (nm)	$\lambda_{\text{émission}}$ (nm)
4',6'-diamidino-2-phénylindole (DAPI)		350	450-490
Isothiocyanate de tetraméthylrhodamine (TRITC)		547	572
Isothiocyanate de fluorescéine (FITC)		495	519
Alexa Fluor 568		579	603

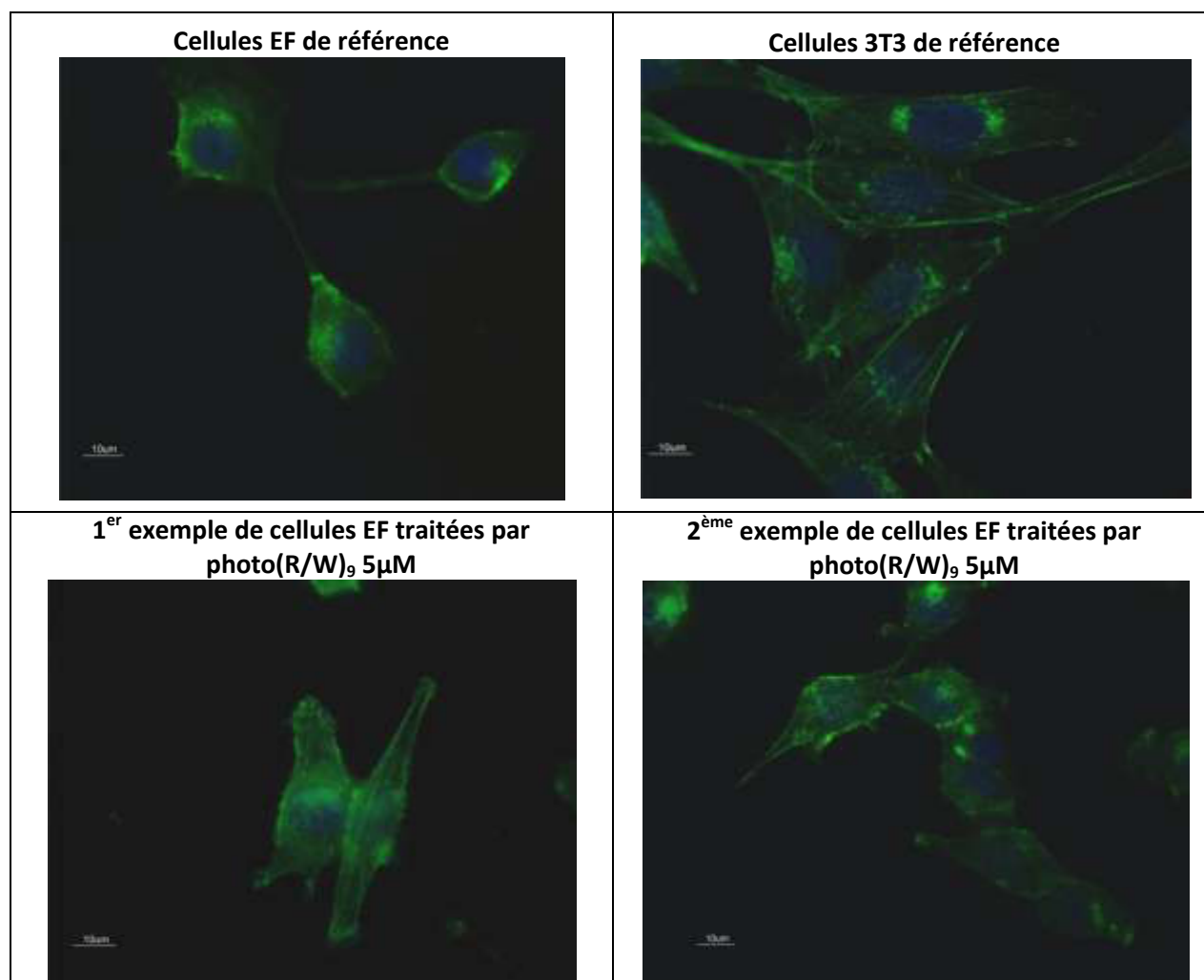
## Résultats : Vers la capture de partenaires par photocross-linking in cellulo

### Expériences réalisées

Après avoir mis au point les conditions de l'expérience d'immunofluorescence (cf. partie matériel et méthodes III.3), l'effet du peptide photo(R/W)<sub>9</sub> sur la morphologie des cellules EF a été observé après 18 h d'incubation à une concentration en peptides de 5  $\mu$ M. Ces cellules ont été comparées avec des cellules soumises à des conditions préalablement étudiées à savoir avec (R/W)<sub>9</sub> réf à 10  $\mu$ M et (R/W)<sub>16</sub> à 10  $\mu$ M.

### Résultats et commentaires

Des exemples d'images représentatives des populations cellulaires observées sont présentés Figure 38.



**Figure 38 : Images de microscopie de fluorescence réalisées pour confirmer l'effet du peptide photo(R/W)<sub>9</sub> sur le remodelage du cytosquelette d'actine dans les cellules EF. Le noyau est marqué avec du DAPI (en bleu) et l'actine F avec de la phalloïdine-FITC (en vert).**

Le travail de thèse réalisé par Diane Delaroche a consisté en grande partie en la caractérisation de l'effet des peptides vecteurs R<sub>9</sub>, (R/W)<sub>9</sub> réf et (R/W)<sub>16</sub> sur le phénotype de cellules EF. Pour l'étude morphologique par immunofluorescence, trois classes de cellules avaient été définies (non tumorale, indéfinie, tumorale) afin de quantifier la proportion de cellules EF de la population traitée par un



peptide ((R/W)<sub>9</sub> réf, (R/W)<sub>16</sub> ou R<sub>9</sub>) présentant un remodelage de leur cytosquelette d'actine et un retour de fibres de stress (cf. partie présentation du projet de recherche). En effet, dans toute population cellulaire, il existe une importante hétérogénéité et il n'est donc pas facile d'affirmer qu'il y a un changement de morphologie dans la mesure où celui-ci ne sera pas visible sur la totalité des cellules. Pour nous il n'était cependant pas question de refaire ce travail extrêmement chronophage car nous souhaitions seulement confirmer que l'introduction de la sonde photoactivable dans la séquence du peptide vecteur (R/W)<sub>9</sub> n'affectait pas son effet biologique. Ainsi nous nous sommes contentés d'une appréciation qualitative et avons comparé différents échantillons de cellules EF traitées par le peptide photo(R/W)<sub>9</sub> par rapport à des cellules EF non traitées, à des cellules 3T3 ou encore des cellules EF traitées par le peptide (R/W)<sub>9</sub> réf, ou le peptide (R/W)<sub>16</sub>. Ces expériences indiquent que **le peptide photo(R/W)<sub>9</sub> a une activité biologique similaire à celle des peptides vecteurs précédemment étudiés.**

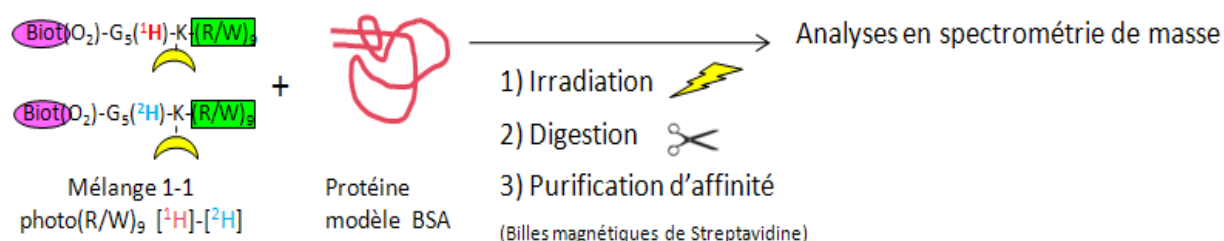
### **Conclusion :**

Les expériences de cytotoxicité, de quantification de l'internalisation du peptide photo(R/W)<sub>9</sub> et d'évaluation par immunofluorescence de son activité de remodelage de l'actine dans les cellules EF nous ont permis de valider biologiquement le peptide photo(R/W)<sub>9</sub> : **ce dernier constitue un modèle adapté** pour aller capturer des partenaires d'interaction *via* la réalisation d'expériences de photocross-linking *in cellulo* afin de mieux appréhender d'un point de vue moléculaire, l'effet observé sur le remodelage de l'actine.

## **II. Validation de la faisabilité de la réaction de photocross-linking *in vitro* et mise au point d'outils**

### **1. Réaction de photocross-linking *in vitro***

En parallèle de la validation biologique du peptide photo(R/W)<sub>9</sub> nous avons testé la réaction de photocross-linking *in vitro* (cf. partie matériel et méthodes III.4). La protéine albumine de sérum de bœuf (BSA) a été notre principal modèle d'étude de la réaction de photocross-linking avec le peptide photo(R/W)<sub>9</sub>. En effet, nous avons pu montrer par des expériences de cross-linking chimique préalables, que (R/W)<sub>9</sub> interagissait avec de multiples zones d'interaction de la BSA. Comme indiqué précédemment (chapitre 5 Partie I.1.), nous avons systématiquement travaillé avec un mélange équimolaire des formes deutérées et non deutérées du peptide photo(R/W)<sub>9</sub> (Figure 39) afin de reconnaître de façon simple les peptides cross-linkés [279].

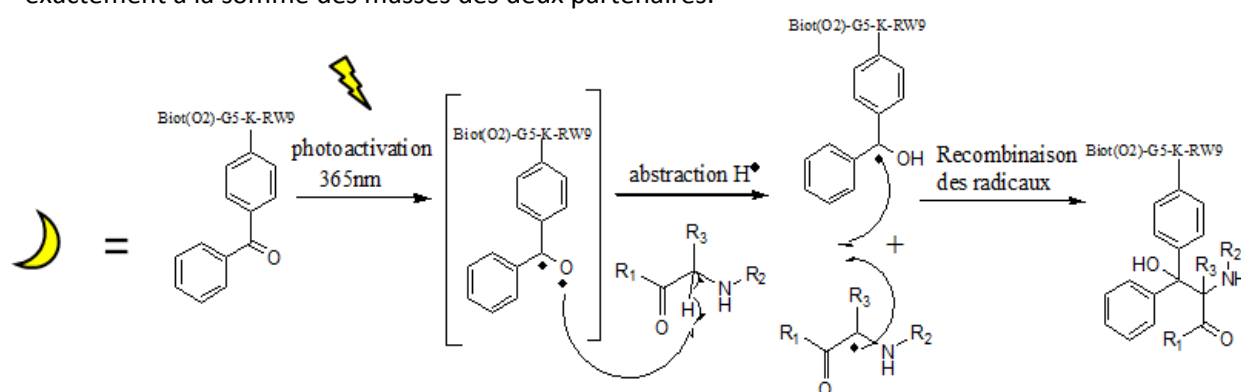


**Figure 39 : Représentation schématique du protocole utilisé pour étudier les peptides cross-linkés issus de la réaction de photocross-linking du peptide photo(R/W)<sub>9</sub> avec la BSA.**



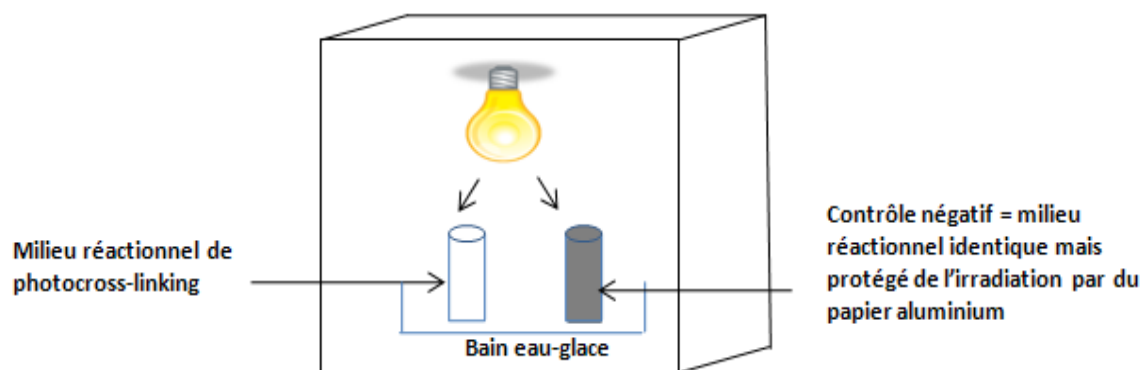
## Résultats : Vers la capture de partenaires par photocross-linking in cellulo

La formation de la liaison covalente entre le peptide photo(R/W)<sub>9</sub> et la protéine partenaire peut impliquer n'importe quelle liaison C-H de n'importe quel acide aminé, à condition que celui-ci soit accessible et donc dans une sphère d'interaction de rayon 0,31 nm autour du carbonyle de la benzophénone. Cette réaction initiée par une irradiation UV à une longueur d'onde proche de 365 nm passe par un mécanisme biradicalaire décrit ci-dessous (Figure 40). La photoirradiation UV entraîne l'activation de la sonde avec la formation d'un biradical C<sup>•</sup>-O<sup>•</sup> sur le carbonyle de la benzophénone. Ensuite, l'O<sup>•</sup> vient arracher un H<sup>•</sup> sur la molécule partenaire conduisant à la formation d'un radical C<sup>•</sup> sur le partenaire et d'une liaison OH sur la benzophénone. Le C<sup>•</sup> restant sur la benzophénone va se recombiner avec le C<sup>•</sup> créé sur le partenaire donnant naissance à une liaison covalente de type C-C. L'espèce photocross-linkée obtenue a une masse moléculaire correspondant exactement à la somme des masses des deux partenaires.



**Figure 40 : Mécanisme de la réaction de photocross-linking entre une sonde de type benzophénone et une protéine.**

Pour la photoirradiation nous avons utilisé le montage artisanal, mais éprouvé, schématisé ci-dessous (Figure 41).



**Figure 41 : Schéma du montage de photoirradiation UV avec une lampe à tungstène (spectre d'émission très large comprenant la longueur d'onde  $\lambda = 365$  nm).**

La lampe utilisée est une lampe à tungstène qui a un spectre d'émission très large comprenant la longueur d'onde d'intérêt, à savoir 365 nm, mais également une zone dans l'infrarouge. Cette lampe va donc avoir tendance à chauffer le milieu réactionnel au cours de l'irradiation. Afin d'éviter l'évaporation de l'échantillon, celui-ci est placé dans un mélange eau/glace pendant toute l'irradiation et à une distance d'environ 5 cm de la lampe (distance optimisée pour un minimum de détérioration de l'échantillon par la chaleur et un rendement de photomarquage optimal).

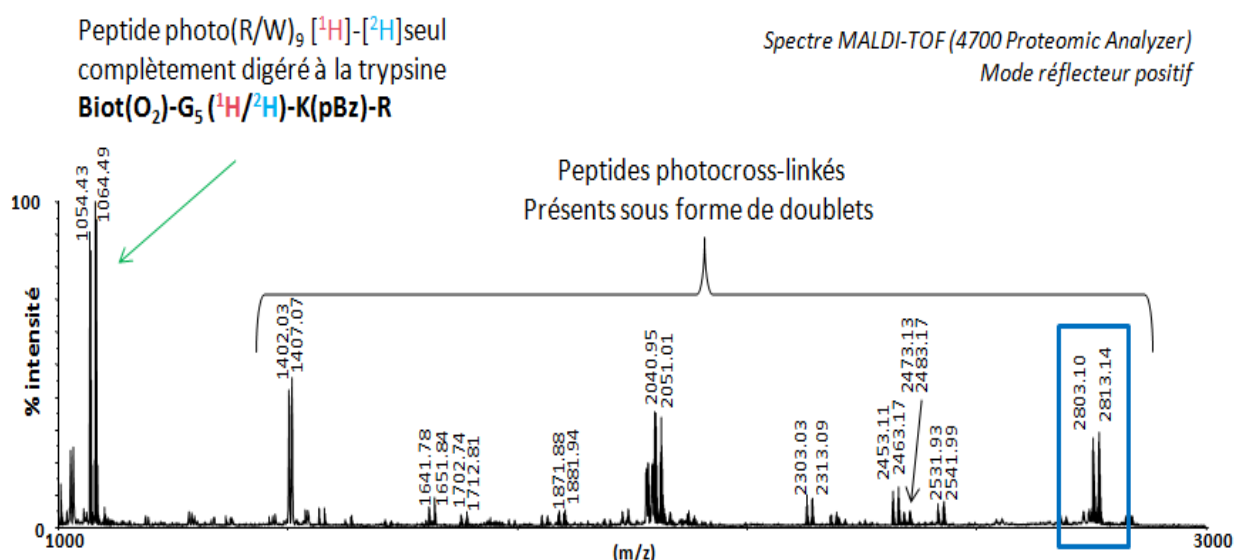
## 2. Analyses des espèces photocross-linkées par spectrométrie de masse MALDI-TOF/TOF et nanoLC-ESI-LTQ-Orbitrap.

Comme indiqué dans la Figure 39, à l'issue de la photoirradiation, une partie du milieu réactionnel est digérée à la trypsine et purifiée sur billes magnétiques de streptavidine de façon à isoler les peptides photocross-linkés. Les peptides piégés sont ensuite élués avec une solution acide. Dans le cas d'une analyse MALDI, l'élution est réalisée avec la matrice CHCA permettant ainsi un dépôt direct sur une plaque MALDI. Dans l'éventualité d'une analyse par nanoLC-ESI-MS/MS, l'élution est réalisée avec une solution d'HCl 0,1M compatible avec une séparation NanoLC.

### a. Analyse en spectrométrie de masse MALDI-TOF/TOF.

#### Analyse MS

L'analyse du milieu réactionnel peptide photo(R/W)<sub>9</sub>-BSA en ratio 1-1 (20 µM-20 µM), obtenu par photoirradiation (30 min) suivie d'une digestion trypsique et d'une purification sur billes magnétiques de streptavidine, donne le spectre de masse présenté Figure 42.



**Figure 42 : Spectre MALDI-TOF (gamme de m/z 1000-3000) obtenu pour le milieu réactionnel de photocross-linking entre photo(R/W)<sub>9</sub> et la BSA à l'issue de la digestion trypsique et de la purification d'affinité.** Matrice utilisée CHCA. Le doublet 2803-2813 correspond aux formes deutérées et non-deutérées d'un peptide photocross-linké. Il est encadré en bleu car nous allons utiliser l'exemple de sa fragmentation dans la suite de ce manuscrit.

Dans ce spectre MALDI-TOF, seuls des ions sous forme de doublets sont observés. Ceci indique que seules les espèces photocross-linkées ou le peptide n'ayant pas réagi ont été enrichis dans ce mélange à l'issue de la purification d'affinité biotine/streptavidine. Le doublet pic de base correspond au peptide photo(R/W)<sub>9</sub> dans sa version la plus courte obtenue après digestion par la trypsine c'est-à-dire au motif Biot-G<sub>5</sub>(<sup>1</sup>H/<sup>2</sup>H)-K(pBz)-R et indique que la réaction de photocross-linking n'est pas totale. Nous observons de nombreux doublets correspondant à des peptides photocross-linkés. Ces doublets ont une intensité allant jusqu'à 50 % de l'intensité du pic de base pour certains, indiquant que la réaction présente un rendement satisfaisant dans ces conditions. La réaction de photomarquage précédemment décrite conduit à la formation d'un complexe covalent dont la masse

## Résultats : Vers la capture de partenaires par photocross-linking in cellulo

est strictement égale à la somme des masses des deux composants, à savoir ici, le motif Biot-G<sub>5</sub>(<sup>1</sup>H/<sup>2</sup>H)-K(pBz)-R du CPP photo(R/W)<sub>9</sub> et le peptide tryptique issu de la BSA. Cette hypothèse de travail nous a permis de générer facilement avec un tableur Excel, la liste des peptides photocross-linkés théoriques et d'attribuer les doublets du spectre MALDI-TOF (Tableau 7). Cela nous a permis d'obtenir notamment une hypothèse robuste sur l'identité du peptide photocross-linké correspondant au doublet 2803-2813. Il s'agirait de la séquence EC!C!HGDLLC!ADDR (avec toutes les cystéines sous forme carbamidométhylée C!) de la BSA covalamment liée au motif Biot-G<sub>5</sub>(<sup>1</sup>H/<sup>2</sup>H)-K(pBz)-R. Ceci a été possible grâce à la bonne précision de la mesure de masse (<20 ppm pour des masses allant jusqu'à 3000 Da) et à une liste de peptides photocross-linkés potentiels relativement restreinte du fait du partenaire protéique unique et connu.

Pour l'échantillon contrôle négatif, à savoir photo(R/W)<sub>9</sub> + BSA (1:1) sans photoirradiation, aucun doublet hormis celui correspondant au peptide photo(R/W)<sub>9</sub> digéré, n'est observé.

**Tableau 7 : Attribution des doublets observés dans le spectre MALDI-TOF présenté (Figure 42).** Les valeurs expérimentales et théoriques des m/z des doublets sont présentées dans les deux premières colonnes. Le delta de masse est donné dans la 3<sup>ème</sup> colonne. Les peptides tryptiques de photo(R/W)<sub>9</sub> et de la BSA sont présentés dans les colonnes suivantes. Le symbole ! indique une modification dynamique, ici la carbamidométhylation des cystéines.

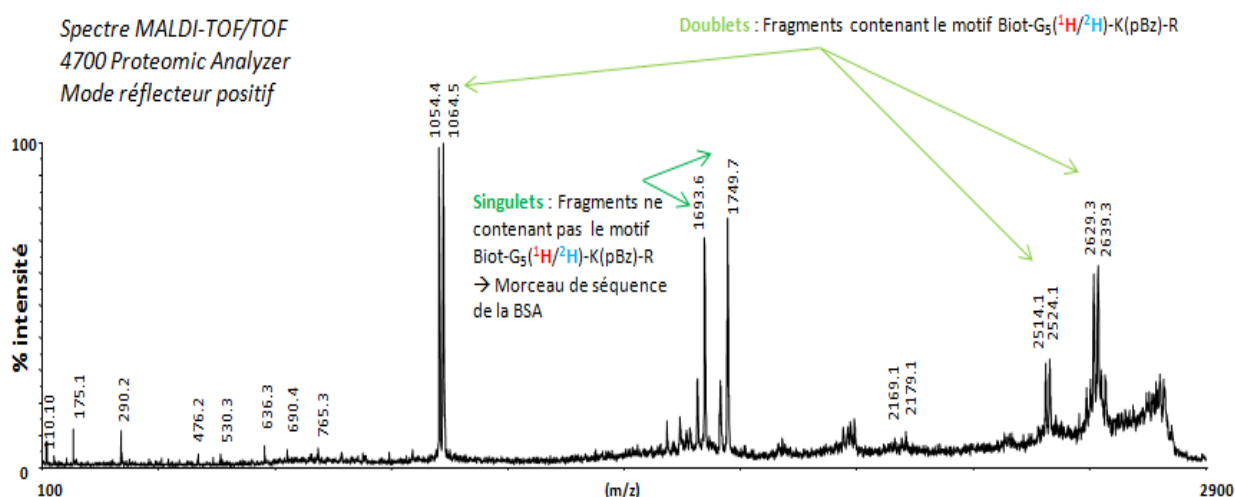
m/z expérimentaux des ions doublets	m/z théoriques des ions doublets	Δmass (ppm)	Fragment photo(R/W) <sub>9</sub>	Fragment BSA	Séquence du peptide tryptique de la BSA
1054,44-1064,49	1064,52-1054,45	19	pep 1-8	non cross-linké	/
1402,54-1407,57 (2+)	1402,05-1407,09 (2+)	14	pep 1-8	267-280	EC!C!HGDLLC!ADDR
1641,78-1651,84	1641,81-1651,87	18	pep 1-9	456-459	VGTR
1702,74-1712,81	1702,77-1712,84	17	pep 1-8	205-209	IETMR
1871,88-1881,94	1871,87-1881,93	5	pep 1-8	562-568	ATEEQLK
2040,95-2051,01	2040,98-2051,05	20	pep 1-8	29-36	SEIAHRFK
2303,03-2313,09	2313,06-2313,13	17	pep 1-8	35-44	FKDLGEEHFK
2453,11-2463,17	2453,13-2463,20	12	pep 1-8	569-580	TVMENFVAFVDK
2473,10-2483,17	2473,13-2483,20	12	pep 1-8	89-100	SLHTLFGDELCLK
2531,93-2541,99	2531,96-2542,03	16	pep 1-8	106-117	ETYGDMADC!CK
2803,08-2813,14	2803,10-2813,17	10	pep 1-8	267-280	EC!C!HGDLLC!ADDR

### Analyse MS/MS

L'étude des spectres de fragmentation des doublets correspondant aux formes deutérées et non-deutérées des peptides photocross-linkés s'est avérée particulièrement intéressante pour caractériser les séquences photomarquées de la BSA et éventuellement, quand cela est possible pour définir le site de photocross-linking. Nous allons détailler l'exemple du doublet 2803-2813 encadré en bleu sur la Figure 42.

Avec le spectromètre de masse MALDI-TOF/TOF (4700 Proteomic Analyzer, Applied Biosystems), la fenêtre de sélection utilisée ne permet pas de sélectionner uniquement un des deux ions du doublet pour la fragmentation. Il est possible dans certains cas, de déplacer cette fenêtre de sélection vers les plus hautes masses (m/z > 2813) ou les plus basses masses (m/z < 2803) pour ne fragmenter qu'un seul des deux ions, mais ceci au détriment de la qualité des spectres MS/MS puisqu'une quantité d'ions non négligeable est perdue lors de cette opération. Nous avons donc choisi de fragmenter simultanément les ions 2803 et 2813 en réglant la fenêtre de sélection sur le rapport m/z 2809 et en choisissant une faible résolution R = 30. Le spectre MS/MS obtenu est présenté Figure 43.

## Résultats : Vers la capture de partenaires par photocross-linking in cellulo

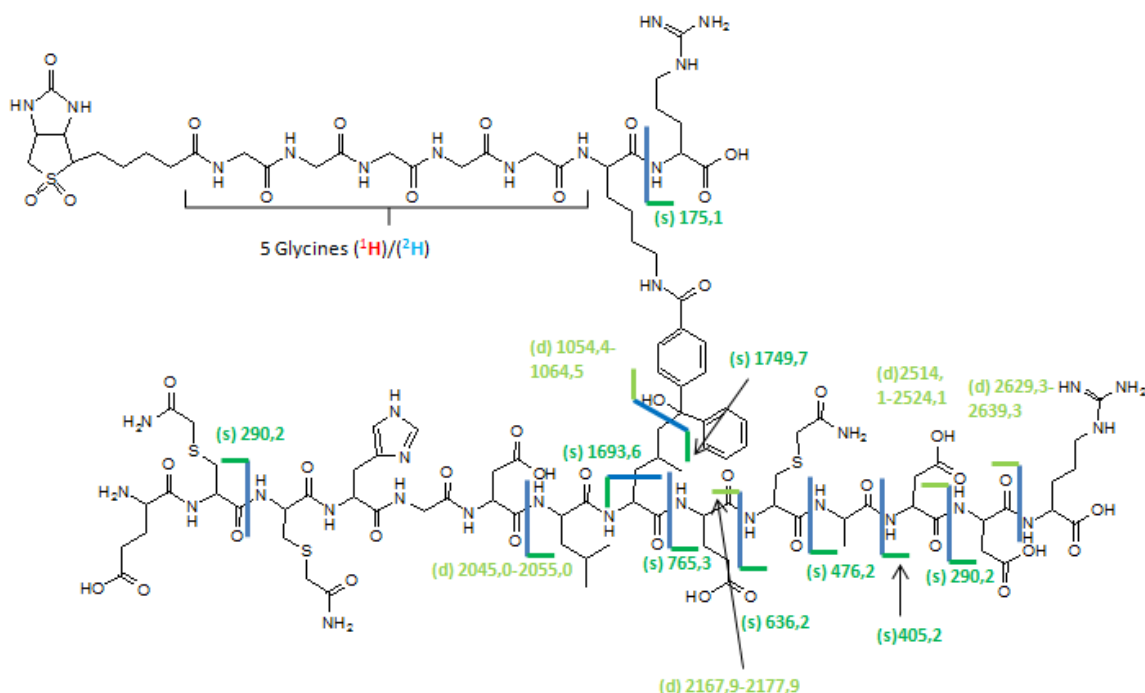


**Figure 43 : Spectre de fragmentation MALDI-TOF/TOF du doublet de peptides photocross-linkés 2803-2813.** Matrice utilisée CHCA.

Sur le spectre de fragmentation nous pouvons immédiatement constater que certains fragments sont sous forme de doublets (d) et d'autres sous forme de singulets (s). Les ions doublets correspondent à des fragments contenant le motif Biot-G<sub>5</sub>(<sup>1</sup>H/<sup>2</sup>H) tandis que les ions singulets ne le contiennent pas. Parmi eux, le pic de base à m/z 1054-1064 correspond à la masse du motif Biot-G<sub>5</sub>(<sup>1</sup>H/<sup>2</sup>H)-K(pBz)-R et résulte de la rupture de la liaison C-C formée entre la benzophénone et la protéine BSA. La rupture préférentielle de cette liaison peut traduire une certaine fragilité de la liaison C-C formée du fait de l'importante contrainte stérique liée à la présence des deux cycles benzéniques de la benzophénone (Figure 44).

L'attribution manuelle des fragments (Figure 43) du spectre représenté, nous a permis de déterminer que, dans le cas du peptide EC!C!HGDLLEC!ADDR, la liaison covalente implique une des deux leucines de la séquence. Le fragment de m/z 1693,6 correspondant à une rupture en le C $\alpha$  et le C $\beta$  de la chaîne latérale d'une leucine est l'élément clé appuyant cette hypothèse.

## Résultats : Vers la capture de partenaires par photocross-linking in cellulo



**Figure 44 : Représentation schématique de la structure potentielle de l'ion du peptide photocross-linké correspondant au doublet 2803-2813.** Attribution manuelle du spectre de fragmentation MALDI-TOF/TOF. Exemple avec l'hypothèse d'une liaison covalente sur la deuxième leucine de la séquence de la BSA. Les barres verticales bleues correspondent aux sites de fragmentation. Les barres horizontales vertes claires sont pour les ions b, et les vertes foncées pour les ions y. (s) pour un ion singulet et (d) pour un doublet.

**Tableau 8 : Attribution des fragments du spectre MALDI-TOF/TOF présenté Figure 43.** Les valeurs des m/z pour les fragments singulets ou doublets sont données dans la première colonne. Le type de fragment b ou y, cross-linké ou non, est indiqué dans la deuxième colonne. Les morceaux de séquence correspondant aux fragments sont indiqués dans les troisième et quatrième colonnes. Le symbole @ donne la position du site de photocross-linking (ici l'hypothèse d'une liaison sur la deuxième leucine est proposée). Le symbole ! indique une modification dynamique ici la carbamidométhylation des cystéines.

m/z fragments observés singulet (s) ou doublet (d)	Type de fragment	Séquence 1 (CPP Photo(R/W) <sub>9</sub> )	Séquence 2 (Fragment du peptide EC!IHGDLL@EC!ADDR de la BSA)
(s) 110,1	ion immonium H	/	/
(s) 175,1	y non cross-linké du CPP ou de la BSA	y--> R	y-->R
(s) 290,2	b non cross-linké ou y non crosslinké	/	b --> E ou y--> DR
(s) 405,2	y non cross-linké	/	y--> DDR
(s) 476,2	y non cross-linké	/	y --> ADDR
(s) 530,3	non attribué	/	/
(s) 636,3	y non cross-linké	/	y--> C!ADDR
(s) 690,4	non attribué	/	/
(s) 765,5	y non cross-linké	/	y--> EC!ADDR
(d) 1054,4-1064,5	non cross-linké	bggggk@r coupe au niveau de la liaison C-C formée par photocross-linking	/
(s) 1693,6	non cross-linké	/	EC!IHGDLL@EC!ADDR coupe entre le Cα et le Cβ de la chaîne latérale du résidu cross-linké
(s) 1749,7	non cross-linké	/	EC!IHGDLL@EC!ADDR coupe au niveau de la liaison C-C formée par photocross-linking
(d) 2044,98-2055,02	y cross-linké	bggggk@r	EC!IHGD
(d) 2169,1-2179,1	b cross-linké	bggggk@r	EC!IHGDLL
(d) 2514,1-2524,1	b cross-linké	bggggk@r	EC!IHGDLLC!AD
(d) 2629,3-2639,3	b cross-linké	bggggk@r	EC!IHGDLLC!ADD

**b. Mise au point d'un outil pour l'interprétation des spectres MS/MS obtenus en NanoLC-ESI-MS/MS : Xlink-Identifier 2.0**

Nous nous sommes ensuite intéressés à l'analyse du même échantillon avec le spectromètre de masse ESI-LTQ-Orbitrap XL de la plateforme. En effet, ce système couplé à une chaîne NanoLC et doté d'une grande résolution est destiné à l'acquisition haut débit de données et capable d'isoler efficacement et automatiquement des espèces pour les fragmenter. Avec ce système nous avons pu obtenir des spectres de fragmentation CID basse énergie et ETD pour tous les peptides photocross-linkés de l'échantillon. Avec ce type d'appareil, seul un des ions du doublet est sélectionné pour être fragmenté.

Nous avons fait appel à **l'outil masstag du logiciel d'acquisition Xcalibur** de manière à fragmenter uniquement les ions visibles sous forme de doublets avec des pics d'intensités proches et espacés de 10,06 Da dans les spectres MS. Ceci permet d'optimiser le temps imparti à la fragmentation en ne fragmentant que les espèces pertinentes. Au-delà de la sélection des espèces d'intérêt à fragmenter, le fait d'avoir à la fois les spectres MS/MS des versions lourdes et légères d'un peptide facilite l'annotation des spectres de fragmentation. Cependant l'attribution manuelle reste fastidieuse même dans le cas de partenaires connus et est donc inenvisageable dans le cas de partenaires d'interactions inconnus présents dans des milieux complexes tels que des lysats cellulaires, par exemple.

Nous avons essayé de contourner le problème en recherchant les espèces photocross-linkées analysées en NanoLC-ESI-MS/MS avec les **moteurs de recherche SEQUEST et MASCOT**. Pour cela, nous avons défini la partie Biot-G<sub>5</sub>(<sup>1</sup>H/<sup>2</sup>H)-K(pBz)-R de ces espèces comme une modification existant sous deux formes, lourde et légère, et pouvant impliquer n'importe lequel des 20 résidus. En effet, la réaction de photomarquage peut entraîner la formation d'une liaison C-C entre la benzophénone et le partenaire sur n'importe quel résidu.

A chaque fois, nous avons pu constater que les ions correspondant aux peptides photocross-linkés, bien que très majoritaires dans le spectre MS après enrichissement sur billes de streptavidine, ne faisaient pas partie de la liste des peptides identifiés avec un FDR < 1 %. Par contre, nous avons pu, par exemple, trouver la paire d'ions 2803-2813 parmi les peptides identifiés avec un FDR > 5 % c'est à dire ceux qui sont classiquement rejetés pour l'identification des protéines (*cf.* tableaux de peptides identifiés pour la BSA en annexes N°1). L'exemple avec l'hypothèse de la liaison sur la leucine est donné Figure 45.

## Résultats : Vers la capture de partenaires par photocross-linking in cellulo

Proteins	Peptides	Search Input	Result Filters	Peptide Confidence	Search Summary									
A2	Sequence	# PSMs	# Proteins	# Protein Groups	Protein Group Accessions	Modifications	IonScore	Exp Value	Charge	MW [Da]	ΔM [ppm]			
1	HPYAVSVLLR	1	2	1	P02769		20	1.4E-001	2	1283.71306	1.85			
2	EIEATLEECAK	1	1	1	P02769	C8(Carbamidomethyl); C10(Carbamidomethyl)	57	8.0E-006	2	1502.61528	0.97			
3	DQRHACIVFDK	1	1	1	P02769	C8(Carbamidomethyl)	38	5.5E-003	2	1554.65471	1.11			
4	ECHGDULEGADDR	4	10	1	P02769	C2(Carbamidomethyl); C3(Carbamidomethyl); C10(Carbamidomethyl)	29	1.6E-003	3	1749.66446	1.06			
5	HPFYAPFLYYANK	3	2	1	P02769		44	2.2E-003	3	1888.92624	-0.33			
6	AKANISVAR	2	2	1	P02769	L2(RW9 p)	18	1.7E+000	3	2055.02304	0.05			
7	ECHGDULEGADDR	4	10	1	P02769	C2(Carbamidomethyl); C3(Carbamidomethyl); L8(RW9p); C10(Carbamidomethyl)	16	6.0E-002	3	2803.79989	1.23			
8	ECHGDULEGADDR	4	10	1	P02769	C2(Carbamidomethyl); C3(Carbamidomethyl); L8(RW9pR0); C10(Carbamidomethyl)	20	7.8E-002	3	2813.15897	-0.08			

**Figure 45 : Peptides correspondant à la protéine BSA (P02769) identifiés lors de la recherche avec Mascot de l'analyse NanoLC-ESI-MS/MS du milieu de photocross-linking photo(R/W)<sub>9</sub>-BSA.** Les deux dernières lignes correspondent aux ions du doublet 2803-2813. Ils sont identifiés avec les modifications notées RW9p ou RW9pR0 correspondant à la partie Biot-G<sub>5</sub>(<sup>1</sup>H/<sup>2</sup>H)-k(pBz)-R sous forme légère ou lourde avec des scores, respectivement, de 16 et de 20 et font ainsi partie des peptides avec un FDR > 5 % (points rouges).

Les moteurs de recherche classiques ne sont donc pas en mesure d'attribuer un score correctement aux peptides photocross-linkés lorsque l'une des parties est définie comme une modification. Ce problème pourrait être lié au fait que plusieurs ions fragments obtenus résultent de fragmentations au sein de cette modification, possibilité qui n'est pas envisagée dans la définition des modifications pour les moteurs de recherche classiques. Ainsi, de nombreux ions ne sont pas attribués, conduisant à des scores trop faibles pour être pris en compte dans les résultats avec un FDR < 1 %.

Pour mener à bien notre stratégie d'identification de partenaires intracellulaires par photocross-linking, il nous faut impérativement être en mesure d'interpréter les spectres de fragmentation des peptides photocross-linkés, rapidement et automatiquement, pour à la fois identifier le partenaire et la zone d'interaction. Nous avons alors fait appel aux bioinformaticiens de l'équipe du Dr. Xiuxia Du, avec laquelle nous avons préalablement travaillé pour l'interprétation des données de cross-linking chimique [280], pour développer une nouvelle version du logiciel Xlink-Identifier capable d'identifier les espèces photocross-linkées.

Ce logiciel a été écrit par Adam Baxter en réalisant les étapes suivantes :

### Identification des peptides photocross-linkés

- **Génération des listes de précurseurs** (peptides photocross-linkés) susceptibles d'être observés à partir des fichiers d'extension .fasta de la protéine BSA et de celui créé pour photo(R/W)<sub>9</sub>.
- Comparaison de la liste des précurseurs fragmentés observés avec la liste théorique des précurseurs générée à l'étape précédente → Identification des peptides photocross-linkés sur la base des ratios m/z des précurseurs (mesure en haute résolution et haute précision en masse).



# Identification du site de photocross-linking

- **Génération de listes de candidats** : Pour chaque peptide photocross-linké identifié il y a autant de site de cross-linking potentiels que de résidus dans le peptide tryptique issu de la protéine qui a été liée covalamment au peptide photoactivable.

Par exemple, pour la forme non-deutérée du peptide photocross-linké comprenant la séquence de la BSA notée ECCHGDLLECADDR (14 résidus) et correspondant à l'ion  $(M+H)^+$  m/z 2803, nous obtenons une liste de 14 candidats pour un spectre MS/MS acquis (Figure 46).

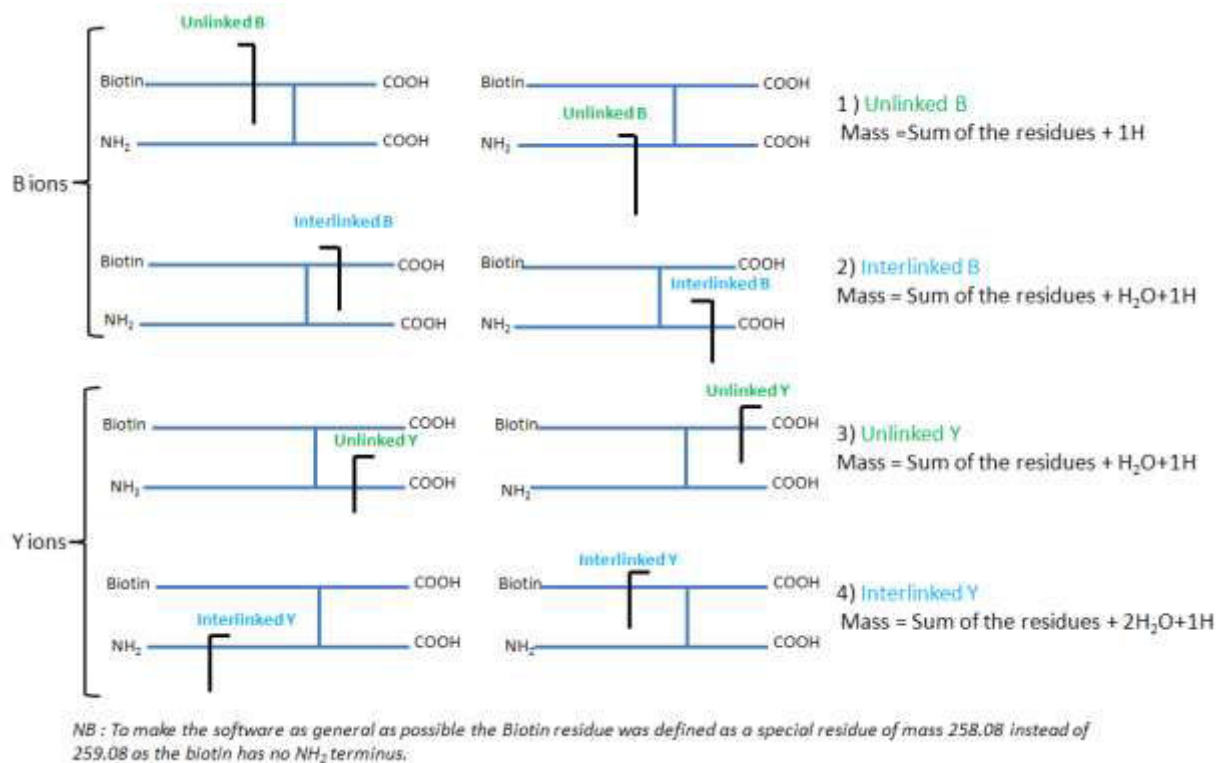
1	2	3	4	5	6	7	8	9	10	11	12	13
XCor	Mass Error	Theoretical Mh	DTA	Linker Config	Linker	Candidate ID	Sequence 1	Decorated 1	Parent 1	Sequence 2	Decorated 2	Parent 2
23.08578	0.34839	2803.0916	BSA-RW9photo-masstag-CID.7637.7637.4	interlinked	photo	1137	bgegggkr	bgegggk@r	CPP	ECCHGDLLECADDR	ECICIHGDLLECIADDR@	BSA
23.26084	0.34839	2803.0916	BSA-RW9photo-masstag-CID.7637.7637.4	interlinked	photo	1131	bgegggkr	bgegggk@r	CPP	ECCHGDLLECADDR	ECICIHGDLLECIADD@R	BSA
23.80041	0.34839	2803.0916	BSA-RW9photo-masstag-CID.7637.7637.4	interlinked	photo	1127	bgegggkr	bgegggk@r	CPP	ECCHGDLLECADDR	ECICIHGDLLECIAD@DR	BSA
25.53622	0.34839	2803.0916	BSA-RW9photo-masstag-CID.7637.7637.4	interlinked	photo	1116	bgegggkr	bgegggk@r	CPP	ECCHGDLLECADDR	ECICIHGDLLECI@ADDR	BSA
25.65215	0.34839	2803.0916	BSA-RW9photo-masstag-CID.7637.7637.4	interlinked	photo	1122	bgegggkr	bgegggk@r	CPP	ECCHGDLLECADDR	ECICIHGDLLECI@DDR	BSA
26.33218	0.34839	2803.0916	BSA-RW9photo-masstag-CID.7637.7637.4	interlinked	photo	1081	bgegggkr	bgegggk@r	CPP	ECCHGDLLECADDR	E@CICIHGDLLECIADDR	BSA
26.8819	0.34839	2803.0916	BSA-RW9photo-masstag-CID.7637.7637.4	interlinked	photo	1084	bgegggkr	bgegggk@r	CPP	ECCHGDLLECADDR	ECI@CICIHGDLLECIADDR	BSA
27.03729	0.34839	2803.0916	BSA-RW9photo-masstag-CID.7637.7637.4	interlinked	photo	1114	bgegggkr	bgegggk@r	CPP	ECCHGDLLECADDR	ECICIHGDLE@CIADDR	BSA
27.42146	0.34839	2803.0916	BSA-RW9photo-masstag-CID.7637.7637.4	interlinked	photo	1087	bgegggkr	bgegggk@r	CPP	ECCHGDLLECADDR	ECICI@HGDLECIADDR	BSA
27.6854	0.34839	2803.0916	BSA-RW9photo-masstag-CID.7637.7637.4	interlinked	photo	1104	bgegggkr	bgegggk@r	CPP	ECCHGDLLECADDR	ECICIHGDL@LECIADDR	BSA
27.69454	0.34839	2803.0916	BSA-RW9photo-masstag-CID.7637.7637.4	interlinked	photo	1100	bgegggkr	bgegggk@r	CPP	ECCHGDLLECADDR	ECICIHGD@LLECIADDR	BSA
27.73261	0.34839	2803.0916	BSA-RW9photo-masstag-CID.7637.7637.4	interlinked	photo	1095	bgegggkr	bgegggk@r	CPP	ECCHGDLLECADDR	ECICIHG@DLLECIADDR	BSA
27.74058	0.34839	2803.0916	BSA-RW9photo-masstag-CID.7637.7637.4	interlinked	photo	1091	bgegggkr	bgegggk@r	CPP	ECCHGDLLECADDR	ECICIH@GDLECIADDR	BSA
32.91448	0.34839	2803.0916	BSA-RW9photo-masstag-CID.7637.7637.4	interlinked	photo	1109	bgegggkr	bgegggk@r	CPP	ECCHGDLLECADDR	ECICIHGDL@ECIADDR	BSA

**Figure 46 : Liste de candidats obtenus avec Xlink-Identifier pour le peptide photocross-linké  $(M+H)^+$  m/z 2803.** Colonne 1 : score calculé par Xlink-Identifier, fonction du nombre de fragments attribués dans le spectre MS/MS ; colonne 2 : erreur en ppm sur la masse du précurseur ; colonne 3 :  $(M+H)^+$  de l'ion précurseur ; colonne 4 : référence du spectre MS/MS acquis et état de charge de l'ion fragmenté (ici spectre MS/MS 7637, ions 4+) ; colonne 7 : numéro du candidat. Les colonnes suivantes font référence aux séquences des deux espèces liées covalamment. Le symbole @ visible dans les colonnes 9 et 12 indique la position du site de photocross-linking.

- **Génération de listes de fragments** pour chacun des candidats susceptibles de correspondre au peptide photocross-linké. Dans le cas des peptides photocross-linkés il faut considérer 4 types de fragments quel que soit le mode de fragmentation (CID ou ETD). A titre d'exemple, ces fragments sont schématisés sur la Figure 47 pour le mode de fragmentation CID, c'est-à-dire avec génération de fragments b et y. Ces fragments, dans la nomenclature utilisée par Xlink-Identifier sont dits "unlinked" si l'ion correspondant est un peptide linéaire et "interlinked" si l'ion correspondant est un peptide branché.



## Résultats : Vers la capture de partenaires par photocross-linking in cellulo



**Figure 47 : Schéma des différents fragments b et y (fragmentation CID) pouvant être obtenus à partir d'un ion précurseur composé de deux peptides liés covalamment par photocross-linking, dont l'un contient une biotine en N-terminal. N.B: la biotine est considérée comme un résidu classique.**

- Comparaison des listes de **fragments des différents candidats** avec les spectres MS/MS acquis pour un peptide photocross-linké. Attribution d'un score à chaque candidat sur la base des pics identifiés. Les différents ions observés et leur attribution sont présentés Figure 48 pour un des candidats (n°1733) correspondant à l'ion précurseur 4<sup>+</sup> de l'ion de (M+H)<sup>+</sup> m/z 2803.

**Spectre MS/MS 8450 candidat N° 1733** Précurseur de M+H<sup>+</sup> =2803,09

First sequence:

Parent - Sequence Id: CPP, Description: Cell penetrating peptide

Sequence: bgggggkr

Link At: 6 Static Mods: DynamicMods:

Second sequence:

Parent - Sequence Id: BSA, Description: gi|30794280|ref|NP\_851335.1| serum albumin precursor [Bos taurus]

Sequence: ECCHGDLLECADDR

Link At: 7

Static Mods:

DynamicMods: {1:!}{2:!}{9:!}

X Correlation Score: 33.845

Sp Score: 92.1559

Matches:

Dta	Theoretical	IonConfig
259.115 : 23.2465	259.07(1 z)	ion: B, {frag seq: alpha, unlinked, alpha: {b}}
290.14 : 12.9648	290.081(1 z)	ion: B, {frag seq: alpha, unlinked, alpha: {EC Dynamic: (!:1)}}
290.14 : 12.9648	290.146(1 z)	ion: Y, {frag seq: alpha, unlinked, alpha: {DR}}
316.104 : 42.9912	316.091(1 z)	ion: B, {frag seq: alpha, unlinked, alpha: {bg}}
318.508 : 6.22408	318.624(2 z)	ion: Y, {frag seq: alpha, unlinked, alpha: {CADDR Dynamic: (!:1)}}
373.099 : 19.1911	373.113(1 z)	ion: B, {frag seq: alpha, unlinked, alpha: {bgg}}
383.461 : 7.08419	383.146(2 z)	ion: Y, {frag seq: alpha, unlinked, alpha: {ECADDR Dynamic: (!:1)}}
405.442 : 5.30276	405.173(1 z)	ion: Y, {frag seq: alpha, unlinked, alpha: {DDR}}
430.157 : 27.041	430.134(1 z)	ion: B, {frag seq: alpha, unlinked, alpha: {bggg}}
476.309 : 12.6626	476.21(1 z)	ion: Y, {frag seq: alpha, unlinked, alpha: {ADDR}}
487.174 : 5.86961	487.156(1 z)	ion: B, {frag seq: alpha, unlinked, alpha: {bgggg}}
544.126 : 4.83827	544.177(1 z)	ion: B, {frag seq: alpha, unlinked, alpha: {bggggg}}
636.383 : 25.117	636.241(1 z)	ion: Y, {frag seq: alpha, unlinked, alpha: {CADDR Dynamic: (!:1)}}
644.27 : 3.3703	644.192(1 z)	ion: B, {frag seq: alpha, unlinked, alpha: {ECCHG Dynamic: (!:2)}}
765.344 : 46.7424	765.284(1 z)	ion: Y, {frag seq: alpha, unlinked, alpha: {ECADDR Dynamic: (!:1)}}
872.598 : 5.99649	872.303(1 z)	ion: B, {frag seq: alpha, unlinked, alpha: {ECCHGDL Dynamic: (!:2)}}
966.548 : 4.44269	966.402(2 z)	ion: Y, {frag seq: alpha, interlinked, alpha: {LECADDR Dynamic: (!:1)}, beta: {bggggkr}}
1019.9 : 8.94061	1019.91(2 z)	ion: B, {frag seq: alpha, interlinked, alpha: {ECCHGDL Dynamic: (!:2)}, beta: {bggggkr}}
1023.09 : 2.41742	1022.94(2 z)	ion: Y, {frag seq: alpha, interlinked, alpha: {LECADDR Dynamic: (!:1)}, beta: {bggggkr}}
1080.82 : 2.51512	1080.46(2 z)	ion: Y, {frag seq: alpha, interlinked, alpha: {DLLECADDR Dynamic: (!:1)}, beta: {bggggkr}}

**Figure 48 : Exemple de fichier candidat répertoriant les pics identifiés dans un spectre MS/MS donné (ici le N°8450) correspondant à un peptide photocross-linké (ici (M+H)<sup>+</sup> m/z 2803).** N.B. : le premier résidu de chaque séquence peptidique est numéroté zéro.

En guise de remarque, on peut noter que le score attribué à un candidat constitue une information intéressante pour l'identification du site de photocross-linking. Elle n'est malgré tout pas suffisante pour trancher sur la position du site de photocross-linking car :

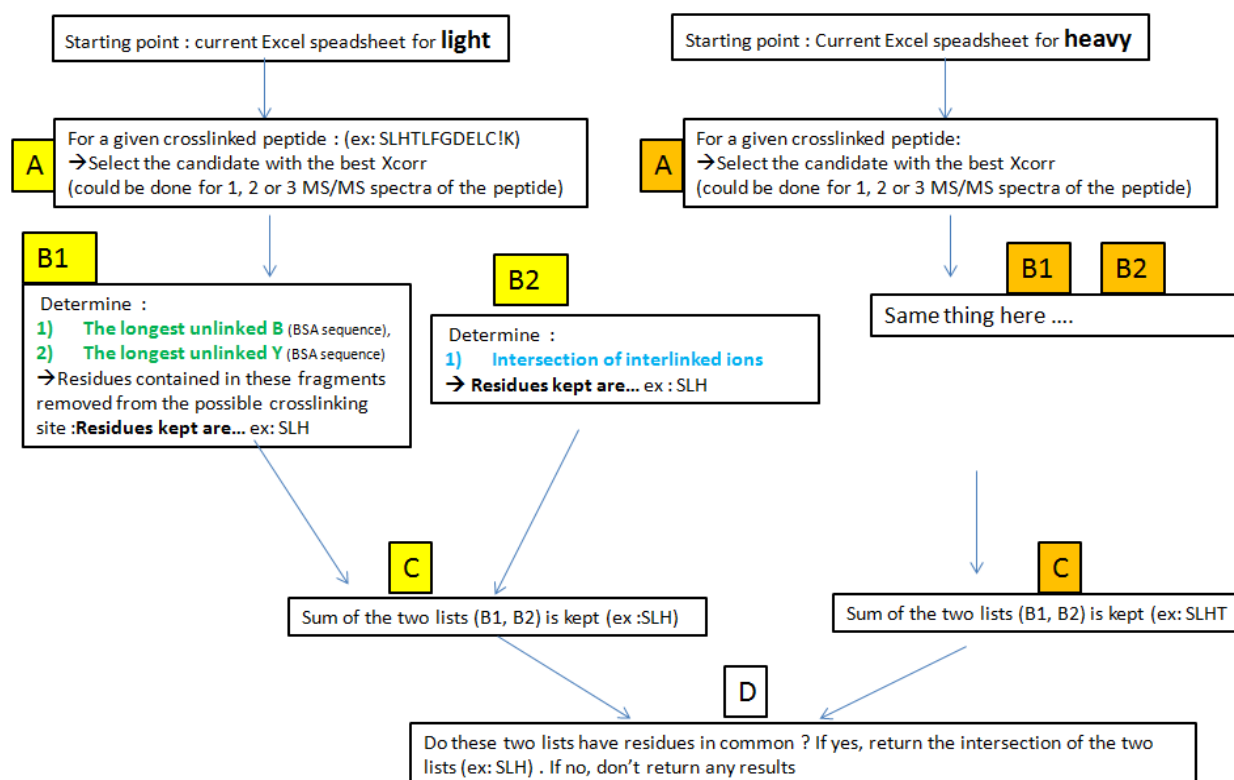
1) Tous les types de fragments sont traités de la même manière : en effet, le spectre d'un ion interlinked avec peu de fragments mais informatifs pourra avoir un score plus faible qu'un spectre comportant de nombreux ions fragments non informatifs sur la localisation du site de photomarquage ;

2) Plusieurs candidats peuvent avoir des scores similaires.

En travaillant sur les listes de fragments, nous avons cherché à établir une procédure pour restreindre à quelques, voire à un seul dans l'idéal, le nombre de résidus potentiellement impliqués dans la réaction de photocross-linking. Cette procédure est schématisée Figure 49.

## Résultats : Vers la capture de partenaires par photocross-linking in cellulo

Only for crosslinked peptides having an MS/MS for both heavy and light version



**Figure 49 : Procédure d'identification du (ou des) résidu(s) potentiellement impliqué(s) dans la réaction de photocross-linking.** Les parties gauche (lettres jaunes) et droite (lettres oranges) de la procédure font respectivement référence au traitement des résultats obtenus pour les versions légères, non-deutérées et pour les versions lourdes, deutérées des peptides photocross-linkés. A l'étape **A**, le candidat qui présente le meilleur score (cf **Figure 46**) est sélectionné. A l'étape **B**, les fragments les plus informatifs (B1 : fragments non cross-linké, B2 : fragments cross-linkés) quant à la position de la liaison sont déterminés et permettent de restreindre le nombre de sites de photocross-linking potentiels. A l'étape **C**, les listes de résidus restants parmi les possibilités c'est-à-dire les résidus potentiellement liés par la benzophénone, sont fusionnées. A l'étape **D**, les listes obtenues pour les versions légères et lourdes des peptides photocross-linkés sont confrontées.

Nos collaborateurs travaillent actuellement à l'écriture du code pour cette procédure ainsi qu'à l'interfaçage du logiciel. Le logiciel, en cours de mise au point, devrait être en mesure d'interpréter à la fois les spectres obtenus en utilisant les modes de fragmentations CID et ETD, ce qui devrait faciliter l'identification des sites de photocross-linking. En effet, les méthodes de fragmentation CID et ETD reposent sur des mécanismes bien distincts qui donnent naissance à des fragments également différents et génèrent donc des informations complémentaires.

**c. Etude comparative manuelle des spectres CID, haute et basse énergie, et ETD d'espèces photocross-linkées.**

De manière à pouvoir optimiser l'interprétation des spectres MS/MS par le logiciel Xlink-Identifier, il nous a semblé important d'annoter précisément les spectres MS/MS CID et ETD de différents peptides photocross-linkés afin de mieux utiliser l'information générée dans ces spectres. Il est important de noter ici que la fragmentation d'espèces cross-linkées n'est pas triviale, en particulier parce que ce sont généralement des espèces de haute masse moléculaire. En effet, en première approximation, leur masse correspond à 2 fois la masse d'un peptide tryptique, en moyenne autour de 1000-1500 Da, soit une masse généralement comprise entre 2000 et 3000 Da. La quantité d'énergie nécessaire à la fragmentation d'un ion de cette masse sera plus élevée que pour un peptide classique du fait de la possibilité d'isomérisation de l'énergie interne sur un plus grand nombre de degrés de liberté. Ceci peut être en partie contrebalancé par le fait que ces peptides branchés sont souvent multichargés du fait de la présence de deux extrémités N-terminales et deux extrémités C-terminales. Enfin, la liaison covalente entre les deux peptides apporte une contrainte stérique qui est susceptible d'influencer les différentes voies de fragmentation.

Dans cette partie, nous confronterons les différentes informations obtenues en fragmentant les espèces photocross-linkées en CID haute énergie (spectres acquis en MALDI-TOF/TOF), CID basse énergie et ETD (spectres acquis en ESI-LTQ/Orbitrap).

➤ **Rappels sur les modes de fragmentation CID et ETD**

La méthode de fragmentation la plus communément utilisée en spectrométrie de masse en tandem des protéines et des peptides repose sur des collisions plus ou moins énergétiques avec des molécules de gaz neutres ( $N_2$  ou He). Ce processus d'activation par collision suivi de la fragmentation des ions est désigné par le terme de **collision induced dissociation (CID)**. Deux types de méthodes CID dites de « **haute** » et de « **basse** » énergie sont généralement distingués [281].

**CID basse énergie**

Les méthodes de CID basses énergies, qui restent encore les plus utilisées, sont réalisées dans des pièges à ions ou des quadripôles et font intervenir des **collisions multiples** avec des molécules de gaz neutres (He, Ar,  $N_2$ ) dans la **gamme d'énergie 1-100 eV**. La durée d'activation des ions est jusqu'à trois ordres de grandeur supérieure à celle utilisée en haute énergie [281]. L'énergie interne acquise au cours de ces collisions va se répartir sur tous les degrés de liberté de la molécule, et ce n'est que lorsque l'énergie interne accumulée est suffisante que la molécule va se fragmenter. Ces méthodes permettent de cliver les ions au niveau des liaisons les plus fragiles, c'est à dire les liaisons amides du squelette peptidique, conduisant aux ions b (fragment contenant la région N-terminale du peptide) et y (fragment contenant la région C-terminale du peptide) et vont également entraîner la perte des PTM labiles.

**CID haute énergie**

Les méthodes de **CID haute énergie**, compatibles avec les analyseurs TOF-TOF, font intervenir un **faible nombre de collisions** mais qui sont réalisées avec une **énergie cinétique importante**, de l'ordre du kiloélectronvolts (KeV). Les méthodes de CID haute énergie conduisent à la rupture

## Résultats : Vers la capture de partenaires par photocross-linking in cellulo

efficace, hétérolytique des **liaisons amides** des peptides en donnant des ions b et y, et permettent également de réaliser des **ruptures des chaînes latérales des acides aminés** (ions d, v, w) comme schématisé sur Figure 50 [282].

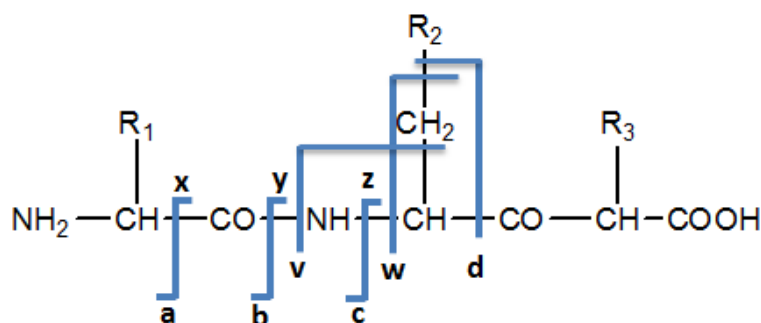


Figure 50 : Représentation des sites de fragmentation d'un peptide selon la nomenclature de Biemann [282].

Les liaisons labiles des modifications post-traductionnelles (PTM) sont généralement rompues. La fragmentation CID haute énergie permet également d'obtenir des ions immoniums résultant de fragmentations multiples.

### ETD

La technique **ETD pour electron transfer dissociation**, développée par Hunt *et al.* [283] est dérivée de la méthode ECD *electron capture dissociation* mise au point par McLafferty *et al.* sur le spectromètre de masse à résonance cyclotronique ionique à transformée de Fourier (FT-ICR) [284]. Elle a l'avantage de pouvoir être conduite sur de nombreux autres types d'appareils (Q-TOF, LTQ par exemple). Dans une expérience ETD, le transfert d'électrons ne se fait pas directement sur l'ion analyte (comme en ECD) mais à partir d'un anion radical. Les anions radicaux sont obtenus à partir d'un composé neutre, typiquement du fluoranthène, capable de capturer un électron thermique produit par ionisation électronique du gaz  $N_2$  (cf. Figure 51).

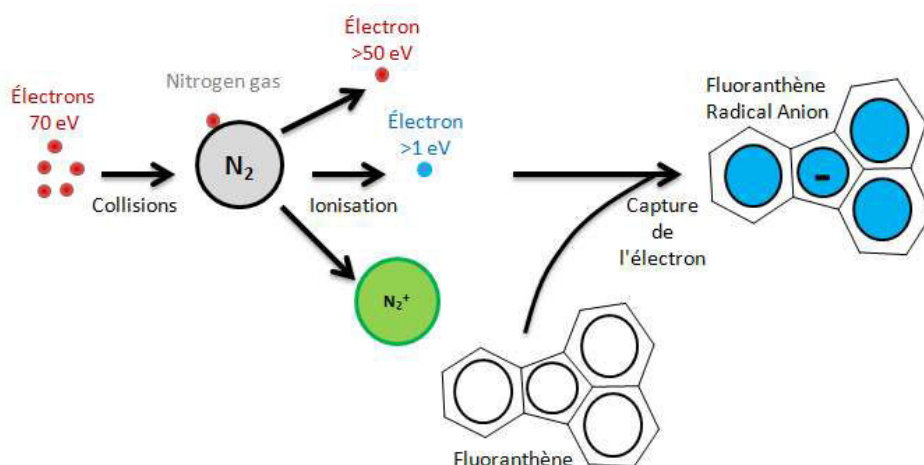


Figure 51 : Production du fluoranthène radical anion dans le volume d'ion, schéma adapté du manuel ETD Thermo.

L'électron va ensuite être transféré rapidement du fluoranthène aux ions précurseurs multichargés sélectionnés. Ce transfert d'électron exothermique (environ 6 eV) va conduire à la fragmentation

## Résultats : Vers la capture de partenaires par photocross-linking in cellulo

rapide, (sans redistribution préalable de l'énergie interne sur tous ses degrés de liberté de l'ion) au niveau du site de capture de l'électron, favorisant la **rupture homolytique des liaisons N-C $\alpha$** , et donner des ions c et z, laissant les chaînes latérales intactes. De plus, les PTM sont généralement conservées [285]. La méthode de fragmentation ETD présente généralement un rendement de fragmentation inférieur à la méthode de CID basse énergie. Cela est lié notamment à la formation d'ions précurseurs à état de charge réduit qui sont des ions dans un état semi-dissocié, dont les fragments sont maintenus ensemble par des liaisons non-covalentes. Pour augmenter le signal des spectres MS/MS il est possible d'utiliser une activation supplémentaire (*supplemental activation*) des ions lors de la fragmentation ETD [286]. Cette activation supplémentaire, de type CID très basse énergie (généralement énergie de collision normalisée de 20%), doit permettre de déstabiliser les liaisons non-covalentes des espèces à état de charge réduit (production d'ions c et z), sans générer de fragmentations supplémentaires.

Pour toutes les raisons précédemment décrites, les deux méthodes de fragmentations (CID et ETD) sont généralement considérées comme des méthodes complémentaires pour l'étude des peptides. Nous avons voulu savoir ce qu'il en était dans le cas de l'étude d'espèces photocross-linkées c'est à dire de peptides « branchés ».

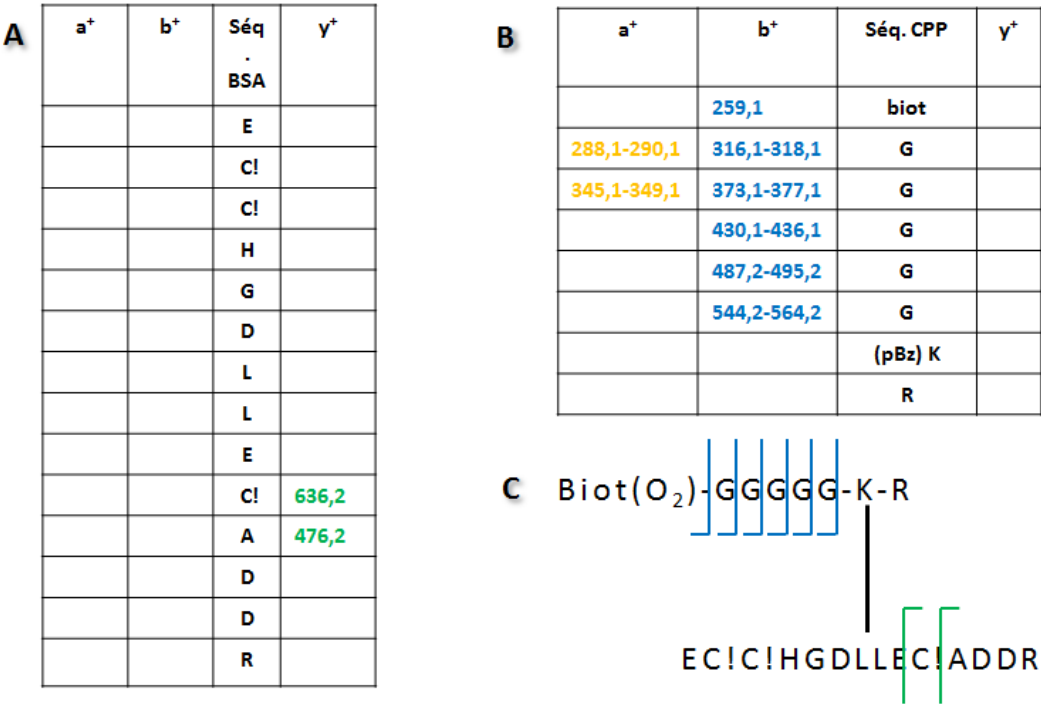
Nous avons cherché à extraire toute l'information pertinente présente sur les spectres de fragmentation de manière à pouvoir, par la suite, optimiser la détermination du site de photocross-linking avec le moteur de recherche Xlink-Identifier.

### ➤ **Annotation des spectres : l'exemple de la paire d'ion (M+H)<sup>+</sup> m/z 2803-2813**

Par soucis de clarté et de concision, les résultats de l'annotation manuelle des spectres de fragmentations sont donnés sous forme de figures : Figure 52 pour les CID basse énergie, Figure 53 pour les spectres ETD et Figure 54 pour les spectres CID haute énergie. Ces figures sont composées de tableaux récapitulant les différents fragments identifiés et de la structure schématique des ions photocross-linkés afin de repérer facilement les sites de rupture. Les spectres correspondant sont présentés en annexes N°2.

*Résultats : Vers la capture de partenaires par photocross-linking in cellulo*

**CID basse énergie**

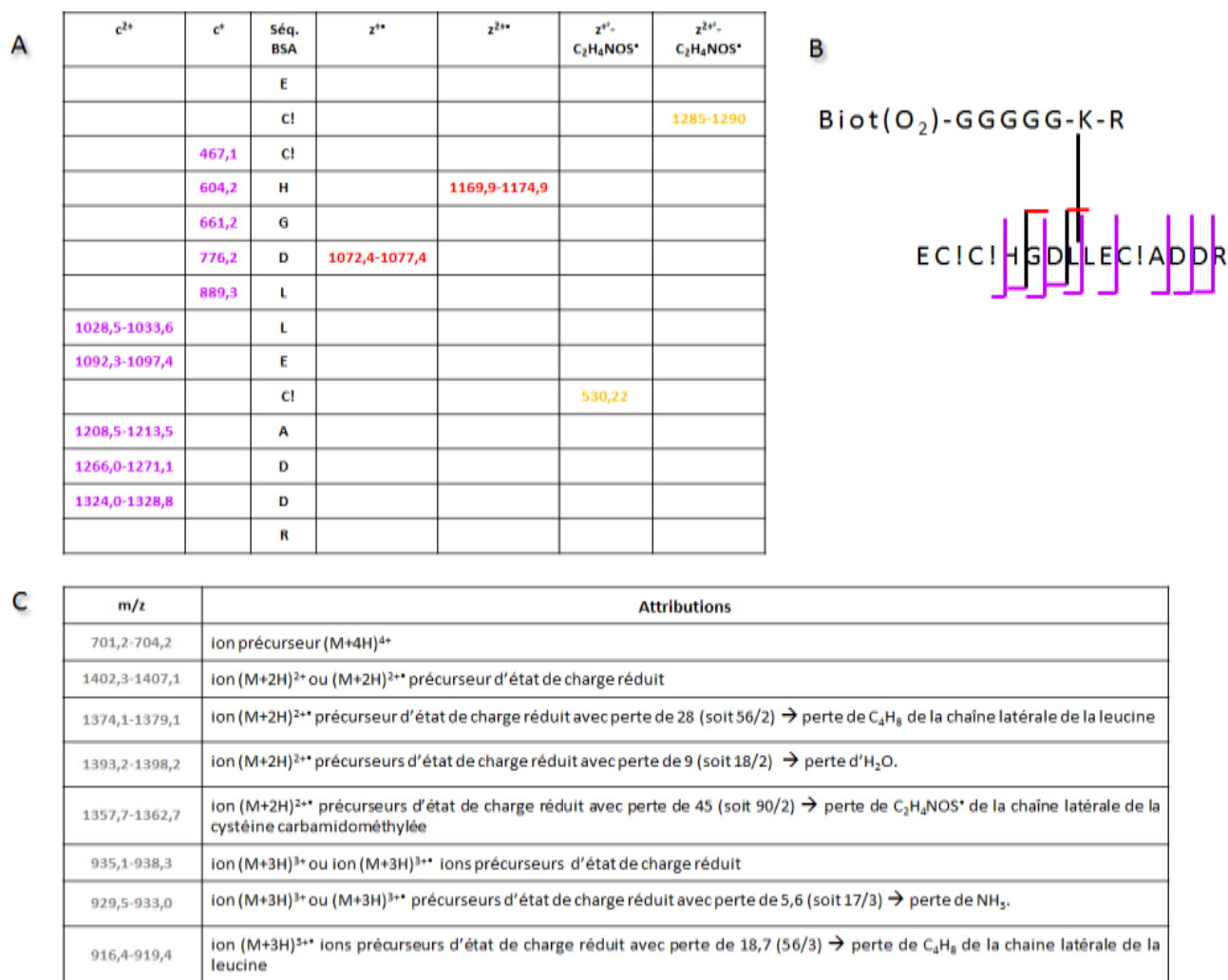


**Figure 52: Récapitulatif des ions observés en CID basse énergie (ESI-LTQ/orbitrap) pour l'ion photocross-linké de masse moléculaire 2803-2813 chargé 4+.** Les parties **A** et **B** de la figure correspondent aux fragments observés dans la séquence du peptide tryptique de la BSA et dans la séquence du peptide tryptique de photo(R/W)<sub>9</sub>. La partie **C** est un schéma récapitulant les sites de coupure sur le peptide photocross-linké.



## Résultats : Vers la capture de partenaires par photocross-linking in cellulo

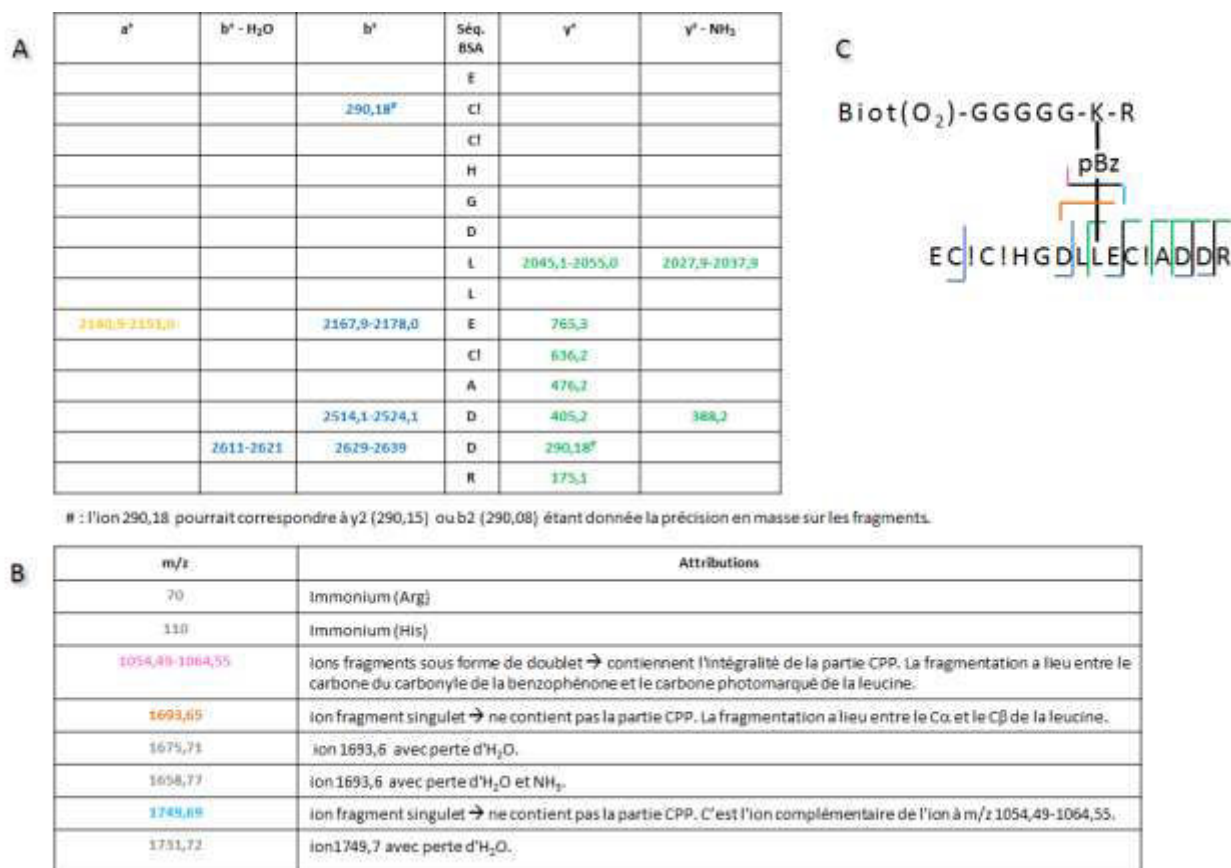
### ETD avec activation supplémentaire (20%)



**Figure 53 : Récapitulatif des ions observés en ETD (avec énergie supplémentaire de 20 %) (ESI-LTQ/orbitrap) pour l'ion photocross-linké de masse moléculaire 2803-2813 chargé 4<sup>+</sup>.** La partie **A** de la figure récapitule les séries ions c et z observés qui sont soit des interlinked (ions doublets) soit des unlinked. La partie **B** est un schéma récapitulant les sites de coupure sur le peptide photocross-linké. La partie **C** liste les ions observés sur le spectre avec une intensité importante. Ces ions correspondent à l'ion précurseur, à l'ion précurseur avec un état de charge réduit ayant subi ou non des pertes de neutres.

## Résultats : Vers la capture de partenaires par photocross-linking in cellulo

### CID haute énergie



**Figure 54 : Récapitulatif des ions observés en CID haute énergie (MALDI-TOF/TOF) pour l'ion photocross-linké de (M+H)<sup>+</sup> 2803-2813 chargé 1<sup>+</sup>.** La partie **A** de la figure correspond aux fragments a, b et y observés. La partie **B** liste les autres ions observés, utiles pour définir le site de photocross-linking. La partie **C** est un schéma récapitulatif des sites de coupure sur le peptide photocross-linké.

Pour faciliter la comparaison, les différentes caractéristiques observées sur les spectres de fragmentation obtenus pour les trois modes, CID haute énergie (MALDI-TOF/TOF), CID basse énergie (ESI-LTQ/Orbitrap) et ETD (ESI-LTQ/Orbitrap) dans le cas de l'analyse de l'ion photocross-linké de (M+H)<sup>+</sup> m/z 2803-2813 sont résumées dans le Tableau 9.

Tableau 9 : Tableau récapitulant les observations et comparaisons réalisées pour les modes de fragmentations CID, haute et basse énergie, et ETD des ions photocross-linkés 2803-2813.

	CID haute énergie (MALDI-TOF/TOF)	CID basse énergie (ESI-LTQ/Orbitrap)	ETD (ESI-LTQ/orbitrap)
<b>Sélection de l'ion précurseur</b>	<ul style="list-style-type: none"> <li>- Fenêtre de sélection large. Dans le cas du doublet d'ions 2803-2813, le doublet est sélectionné.</li> <li>- Intérêt pour l'attribution des fragments : les fragments contenant l'étiquette biot-G5 sont sous forme de doublets les autres sont sous forme de singulets.</li> </ul>	<ul style="list-style-type: none"> <li>- Sélection précise.</li> <li>- Possibilité de fragmenter séparément les ions 2803 et 2813 du doublet 2803-2813.</li> </ul>	<ul style="list-style-type: none"> <li>- Sélection précise.</li> <li>- Possibilité de fragmenter séparément les ions 2803 et 2813 du doublet 2803-2813.</li> </ul>
<b>Intensité de l'ion précurseur</b>	<ul style="list-style-type: none"> <li>- Ion précurseur largement majoritaire dans le cas d'une acquisition en "metastable suppressor off", corrigé et diminué artificiellement sur les spectres pour "metastable suppressor on".</li> </ul>	Non observé.	<ul style="list-style-type: none"> <li>- Ion précurseur majoritaire, accompagné de ses états de charge réduits, malgré l'utilisation d'une activation supplémentaire avec une énergie de collision normalisée de 20 %.</li> </ul>
<b>Pertes de neutres</b>	<ul style="list-style-type: none"> <li>- Pertes d'H<sub>2</sub>O et NH<sub>3</sub> observées sur les fragments.</li> </ul>	Aucune observée.	<ul style="list-style-type: none"> <li>- Pertes intenses observées sur les ions précurseurs à état de charge réduit et sur les ions z cations radicaux (pertes d'H<sub>2</sub>O, de NH<sub>3</sub>, pertes de C<sub>2</sub>H<sub>4</sub>NOS<sup>+</sup> sur Cl<sub>2</sub> de C<sub>4</sub>H<sub>8</sub><sup>+</sup> sur L).</li> </ul>
<b>Fragmentation du squelette peptidique</b>	<ul style="list-style-type: none"> <li>- Séries y et b conduisant à une identification du partenaire mais caractérisation incomplète.</li> </ul>	<ul style="list-style-type: none"> <li>- Peu d'informations structurales, peu d'ions fragment y et b.</li> <li>- Fragments caractéristiques de l'étiquette biot-G5 → ions diagnostiques des complexes cross-linkés (m/z 259, 316, 373, 430, 487, 544.)</li> </ul>	<ul style="list-style-type: none"> <li>- Séries c et z conduisant à une identification du partenaire et une caractérisation complète avec détermination du résidu photomarqué.</li> </ul>
<b>Fragmentations secondaires</b>	<ul style="list-style-type: none"> <li>- Ions immoniums observés.</li> </ul>	Non observées.	Non observées.
<b>Types d'ions fragments</b>	<ul style="list-style-type: none"> <li>- Fragmentation majoritaire au niveau de la liaison C-C formée entre la benzophénone et le peptide de la BSA (gène stérique → fragilité de la liaison).</li> <li>- Fragmentation au niveau de la chaîne latérale du résidu photomarqué → ions complémentaires.</li> <li>→ caractérisation fine possible au niveau de l'atome photomarqué, intérêt pour les études structurales menées par photomarquage.</li> <li>- Unlinked et interlinked pour les ions bet y.</li> </ul>	<ul style="list-style-type: none"> <li>- Fragmentations au niveau du résidu photomarqué non observées.</li> <li>- Quelques unlinked y.</li> </ul>	<ul style="list-style-type: none"> <li>- Fragmentations au niveau du résidu photomarqué non observées.</li> <li>- Unlinked et interlinked pour les ions c et z.</li> </ul>

## **Conclusion**

Les spectres CID basse énergie pour les ions  $(M+H)^+$   $m/z$  2803-2813 se sont avérés être peu informatifs avec seulement la présence de fragments signatures de l'étiquette Biot( $O_2$ )-GGGGG. Ces spectres ont été acquis sur des précurseurs avec un état de charge  $4^+$  qui est l'état de charge majoritaire. En effet, les peptides photocross-linkés obtenus à l'issue de la digestion enzymatique présentent deux extrémités N-terminales et deux extrémités C-terminales soit, au minimum, autant de possibilité de charges positives. Ces ions sont également un peu plus gros que la moyenne et présentent donc davantage de degrés de liberté pour répartir leur énergie interne. Il est ainsi possible que la fragmentation CID basse énergie, qui est classiquement utilisée pour des peptides plus petits avec des états de charges plus bas, soit peu efficace dans le cas d'ions 4 fois chargés comme c'est le cas ici.

Au contraire, le mode ETD est connu pour être plus efficace pour les plus hauts états de charges [287] et cela se confirme pour cet exemple. En effet avec le mode ETD, des séries de fragments c et z ont été obtenues et permettent de déterminer la position du site de photo-crosslinking: sur la deuxième leucine de la séquence EC!CHGDLLEC!ADDR.

Les spectres de fragmentation en CID haute énergie sont également informatifs dans le cas des ions  $(M+H)^+$   $m/z$  2803-2813 avec notamment la précision obtenue grâce la rupture entre le  $C\alpha$  et le  $C\beta$  de la chaîne latérale du résidu photomarcué. De plus, il a été possible avec ce mode de fragmentation de restreindre les possibilités pour le site de photocross-linking aux deux leucines de la séquence. Cependant, pour d'autres peptides photocross-linkés étudiés le CID haute énergie s'est avéré être un peu moins informatif avec un nombre plus faible de fragments obtenus.

De manière générale, c'est le **mode de fragmentation ETD qui semble être le plus adapté pour l'étude des peptides photocross-linkés**. Nous allons faire en sorte d'intégrer au mieux les informations obtenues avec l'attribution manuelle des spectres pour optimiser la recherche des peptides photocross-linkés avec le logiciel Xlink-Identifier 2.0.

### **III. Expériences de photocross-linking *in cellulo***

A l'issue de ces études préliminaires ayant permis d'une part de valider biologiquement le peptide photo(R/W)<sub>9</sub> et d'autre part de confirmer la faisabilité de la réaction de photocross-linking, nous sommes passés aux expériences de photocross-linking *in cellulo*.

Les expériences de photocross-linking *in cellulo* ont été menées sur les cellules EF sur lesquelles l'effet du peptide photo(R/W)<sub>9</sub> a été observé.

#### **1. Evaluation de la cytotoxicité de l'irradiation**

Avant de réaliser l'irradiation de ces cellules en présence de peptides nous nous sommes assurés que les cellules EF supportaient correctement l'irradiation et avons donc testé la cytotoxicité du traitement UV avec le kit CCK-8 déjà utilisé pour évaluer la cytotoxicité des peptides. Aucune cytotoxicité n'a été décelée pour 30 min d'irradiation (suivi de la viabilité cellulaire à 2 h, 3h30 et 5 h).

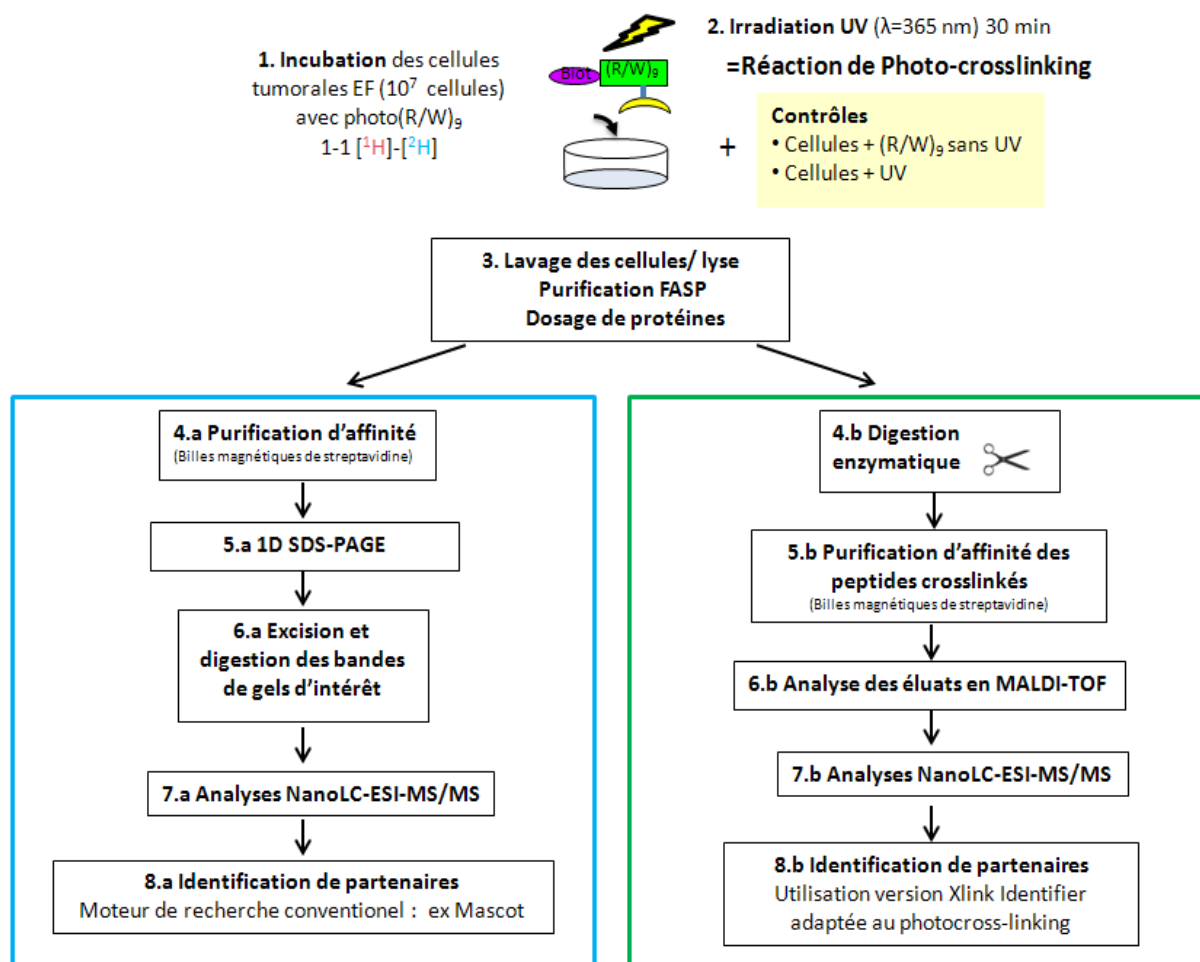
#### **2. Expériences de photocross-linking *in cellulo***

Pour la présentation de cette partie, j'ai choisi de partir du **schéma de la stratégie initialement imaginée**. Les choses ne s'étant pas déroulées comme prévu, nous verrons à quels niveaux de la stratégie nous avons rencontré les principales difficultés et comment nous avons cherché à les surmonter.

##### **a. Présentation et test de la stratégie initiale**

Au départ, nous avons conçu la stratégie d'identification des partenaires intracellulaires ou membranaires du peptide (R/W)<sub>9</sub> présentée schématiquement Figure 55.

## Résultats : Vers la capture de partenaires par photocross-linking in cellulo



**Figure 55 : Stratégie initiale (N°1) envisagée pour l'identification des partenaires d'interaction (intracellulaires et/ou membranaires) du peptide (R/W)<sub>9</sub> par photocross-linking.** Le protocole correspondant est présenté dans la partie matériel et méthodes III.5.

A l'issue de l'incubation des cellules EF avec le peptide photo(R/W)<sub>9</sub> en mélange équimolaire des formes deutérées et non-deutérées, les boîtes de cellules sont placées dans le montage d'irradiation décrit précédemment (cf. Figure 41). A la fin de l'irradiation les cellules sont lavées afin d'éliminer le peptide photo(R/W)<sub>9</sub> non internalisé, puis lysées avec un tampon de lyse stringent permettant de solubiliser la quasi-totalité du matériel cellulaire. Ensuite **deux approches complémentaires**, l'une (à gauche notée branche a.) faisant appel à la purification de complexes entiers et l'autre (à droite notée branche b.) reposant sur la purification de peptides photocross-linkés obtenus après digestion, doivent aboutir à l'identification de partenaires.

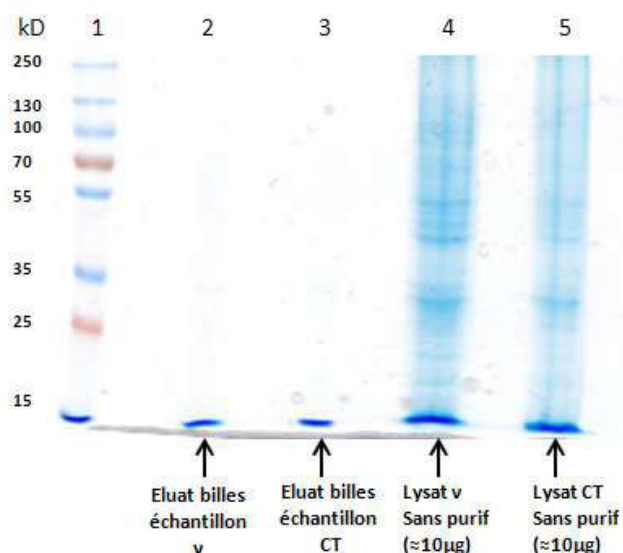
Dans la branche a., l'identification des partenaires peut se faire avec des moteurs de recherche conventionnels dans la mesure où tous les peptides tryptiques d'une (des) protéine(s) partenaire(s) photocross-linkée(s) avec le peptide peuvent être analysés en spectrométrie de masse. Dans la branche de droite, seuls les peptides photocross-linkés sont présents à l'issue de la purification par affinité et l'identification des partenaires devra donc passer par l'utilisation du logiciel Xlink-Identifier développé avec l'équipe du Dr. X. Du et mentionné précédemment.

## Résultats : Vers la capture de partenaires par photocross-linking in cellulo

### Résultats :

Etant donné que les deux branches de la stratégie initiale ne conduisent pas aux mêmes résultats (peptides photocross-linkés *versus* peptides tryptiques de la protéine capturée entière), il convient de présenter successivement les résultats obtenus pour les deux branches.

**Branche a** : Ci-dessous (Figure 56) est présenté le gel 1D SDS-PAGE obtenu à l'étape 5.a (cf. Figure 55).



**Figure 56: Gel 1D SDS-PAGE coloré au bleu de Coomassie.** piste 1 : marqueurs de poids moléculaires, piste 2 : cellules photoirradiées en présence de peptide photo(R/W)<sub>9</sub> (Echantillon v) et purifiées sur billes ; piste 3 : cellules photoirradiées en l'absence de peptide photo(R/W)<sub>9</sub> (Echantillon CT) et purifiées sur billes ; pistes 4 et 5 : environ 10 µg de lysats des échantillons v (irradiés) et CT (contrôle) respectivement.

Le gel coloré au bleu de Coomassie ne présente **aucune bande pour les pistes 2 et 3** correspondant aux milieux purifiés sur billes. Cela était attendu pour le contrôle (noté CT et déposé piste 3) mais pas pour l'échantillon de lysat issu de cellules photoirradiées en présence de peptide photo(R/W)<sub>9</sub> (noté v et déposé piste 2). En effet, les complexes photocross-linkés formés sont en principe purifiés sur billes magnétiques de streptavidine avant d'être élués et migrés sur gel.

En l'absence de bandes d'intérêt, le protocole n'a pas été poursuivi au delà de l'étape 5.

Trois hypothèses principales peuvent être avancées pour expliquer l'absence de bandes de protéines:

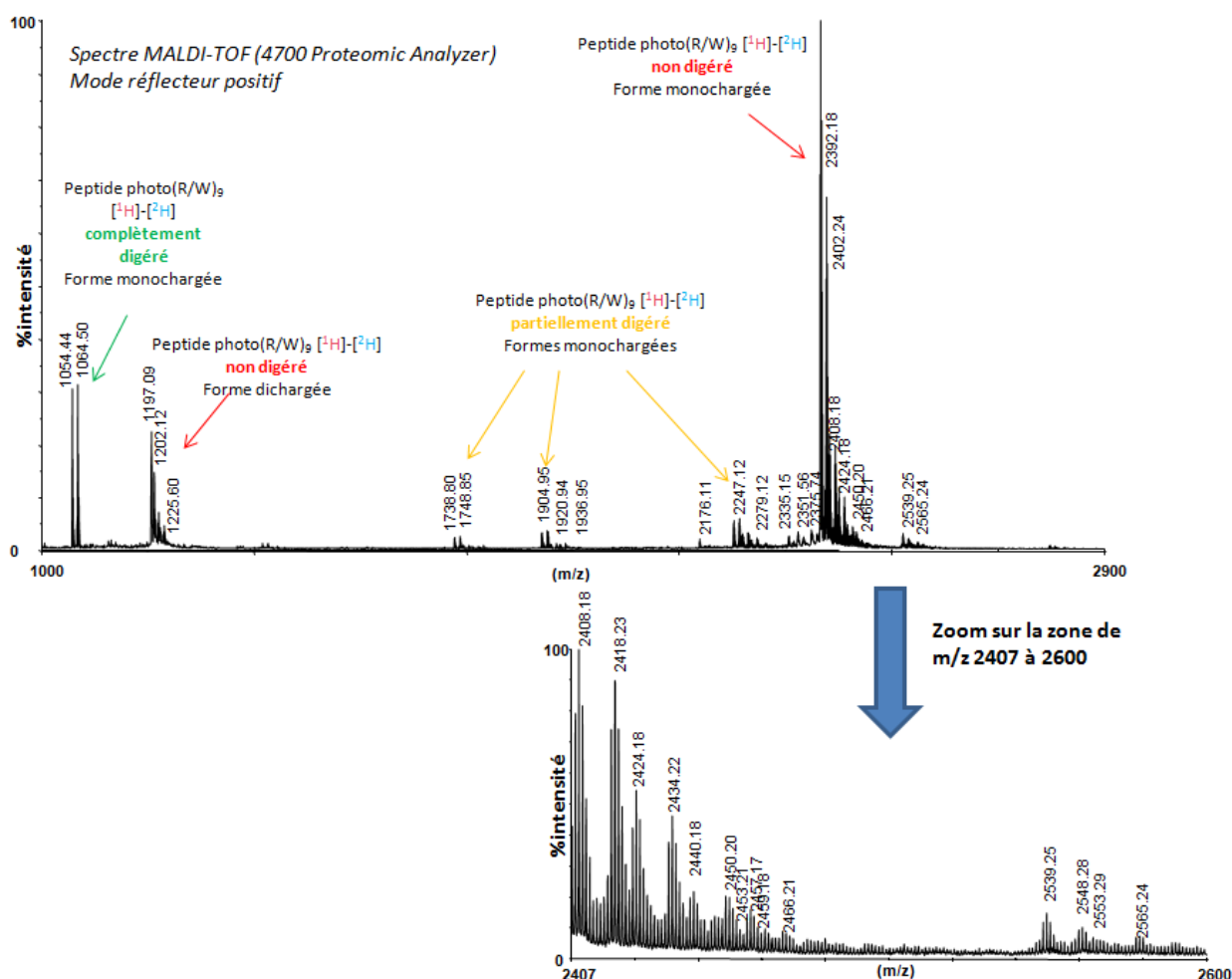
- Un **faible rendement de photocross-linking** et donc une quantité de complexes photocross-linkés trop faible pour donner lieu à des bandes visibles sur le gel.
- Une **grande multiplicité de partenaires d'interaction** se liant avec une affinité moyenne à faible au peptide photo(R/W)<sub>9</sub>.
- Une **mauvaise capture des complexes photocross-linkés entiers** sur billes magnétiques de streptavidine (étiquette biotine portée par le peptide photo(R/W)<sub>9</sub> masquée lorsque ce dernier est photocross-linké à une protéine de taille importante). Il est intéressant de noter que ce problème de purification des complexes entiers formés entre un peptide biotinylé et



## Résultats : Vers la capture de partenaires par photocross-linking in cellulo

une protéine partenaire de masse plus importante a déjà été observé au sein du laboratoire pour d'autres types d'applications en utilisant ces mêmes billes magnétiques greffées avec de la streptavidine.

**Branche b** : Dans le cas de la branche b, l'analyse des éluats correspondant aux milieux réactionnels digérés puis purifiés sur billes magnétiques de streptavidine est réalisée en MALDI-TOF (étape 6.b). Les résultats sont très reproductibles et donnent lieu à des spectres de masse tel que celui présenté Figure 57.



**Figure 57 : Spectre MALDI-TOF obtenu pour l'éluat correspondant au milieu réactionnel v digéré et purifié sur billes. Matrice utilisée CHCA.**

L'interprétation du spectre de masse Figure 57, nous donne les informations suivantes :

- La présence de doublets indique la présence du peptide photo(R/W)<sub>9</sub> et/ou d'espèces photocross-linkées.

Les doublets d'intensités les plus élevées correspondent au peptide photo(R/W)<sub>9</sub> entier (m/z 2392-2402 (1+), 1197-1202(2+)), partiellement digéré (m/z 1720-1730, 1894-1904, 2237-2247) ou complètement digéré (m/z 1054-1064) et non à des espèces photocross-linkées. Il semblerait que les détergents utilisés pour la lyse cellulaire n'aient pas été complètement éliminés, conduisant à une

## **Résultats : Vers la capture de partenaires par photocross-linking in cellulo**

digestion incomplète du peptide photo(R/W)<sub>9</sub> et donc vraisemblablement aussi des protéines. Ceci pourrait expliquer le faible nombre de doublets de plus haut m/z et leur faible intensité. Quelques doublets de moindre intensité et de m/z supérieurs à ceux du peptide photo(R/W)<sub>9</sub> entier sont également visibles (cf. zoom Figure 57). Il pourrait s'agir d'espèces photocross-linkées ou bien d'adduits du peptide photo(R/W)<sub>9</sub> avec des cations. Etant donné leur faible intensité et la proximité du doublet 2393-2402 correspondant au peptide photo(R/W)<sub>9</sub> entier, la largeur de la fenêtre de sélection de l'ion précurseur du MALDI-TOF/TOF est trop importante pour sélectionner correctement ces doublets d'intérêt et obtenir des informations sur leur séquence.

**Conclusion :** Les résultats obtenus pour l'expérience réalisée selon la stratégie initiale, nous conduisent à nous poser les questions suivantes:

- *Y-a-t-il bien formation de complexes photocross-linkés dans les cellules et avec quel rendement ?*
- *Est-il possible de purifier ces complexes photocross-linkés entiers sur billes ?*

Ils nous incitent également à optimiser certaines étapes du protocole à savoir l'élimination des détergents avant la digestion et la quantité d'enzyme à utiliser.

- *L'identification de partenaires d'interactions sera-t-elle possible avec la branche b à l'issue de l'optimisation des conditions de digestion ?*

C'est sur la base de ces premiers résultats et des questions qu'ils ont suscitées que nous avons décidé de réaliser des expériences complémentaires.

### **b. Expériences complémentaires**

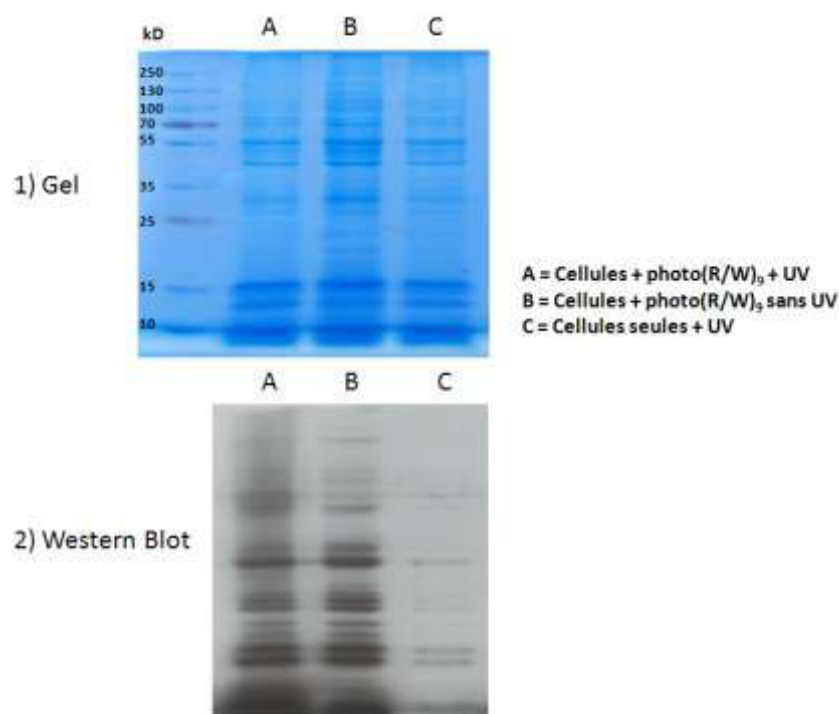
- *Y-a-t-il bien formation de complexes photocross-linkés dans les cellules et avec quel rendement?*

Pour répondre à cette question nous avons réalisé une nouvelle expérience de photoirradiation sur cellules, cellules qui ont ensuite été lysées. Trois échantillons distincts ont ainsi été obtenus :

- Un lysat de cellules incubées en présence de photo(R/W)<sub>9</sub> et photoirradiées (A) = expérience
- Un lysat de cellules incubées en présence de photo(R/W)<sub>9</sub> mais non-irradiées (B) = contrôle
- Un lysat de cellules irradiées en l'absence de peptide photo(R/W)<sub>9</sub> (C) = contrôle

Une quantité similaire de chacun de ces trois milieux a été migrée sur trois pistes de gel SDS-PAGE 1D (pistes A, B et C) en duplicat. L'un des gels a été coloré au bleu de Coomassie, l'autre a été transféré sur membrane de nitrocellulose pour effectuer un western blot en utilisant de la streptavidine couplée à une enzyme HRP (*horseradish peroxidase*) capable de reconnaître la biotine portée par le peptide photo(R/W)<sub>9</sub> avec une grande affinité et ainsi visualiser les espèces photocross-linkées.

## Résultats : Vers la capture de partenaires par photocross-linking in cellulo



**Figure 58 :** Image de 1) gel 1D SDS-PAGE coloré au bleu de Coomassie et de 2) western-blot réalisé avec de la streptavidine HRP. Piste A : cellules + photo(R/W)<sub>9</sub> + UV ; piste B : cellules + photo(R/W)<sub>9</sub> sans UV ; piste C : cellules + UV.

### Résultats :

Nous constatons sur le gel 1D SDS-PAGE (Figure 58) que les quantités de lysats que nous avons souhaitées déposer, à savoir 75 µg sont relativement similaires, aux erreurs de dosage et de pipetage près (légèrement plus dans la piste B).

Sur le western blot (Figure 58), nous observons un nombre important de bandes pour la piste A. La grande majorité de ces bandes sont également visibles pour la piste B avec une intensité similaire. Par contre, ces bandes sont très faiblement visibles pour la piste C.

En principe, les bandes observées correspondent à des espèces photocross-linkées contenant une biotine, les complexes non cross-linkés étant censés être dissociés par le SDS. La présence de ce type de bandes était donc attendue pour la piste A. Le nombre important des bandes visibles suggère l'existence d'un grand nombre de partenaires d'interactions. Cependant pour la piste B, correspondant aux cellules incubées en présence de peptide photo(R/W)<sub>9</sub> mais non irradiées, nous ne devrions observer qu'une seule bande située en bas du gel et correspondant au peptide photo(R/W)<sub>9</sub> seul. Toutes les bandes situées au dessus du front de migration ne peuvent pas correspondre à des espèces photocross-linkées et doivent donc être dues à des **associations non-covalentes de protéines avec le peptide photo(R/W)<sub>9</sub> qui sont maintenues malgré le caractère dénaturant du gel 1D-SDS PAGE**. Ces associations sont donc particulièrement fortes.

La présence de bandes de faibles intensités dans la piste C peut, elle, être expliquée par une légère contamination de la piste C par la piste B ou par une reconnaissance non spécifique de la streptavidine-HRP.

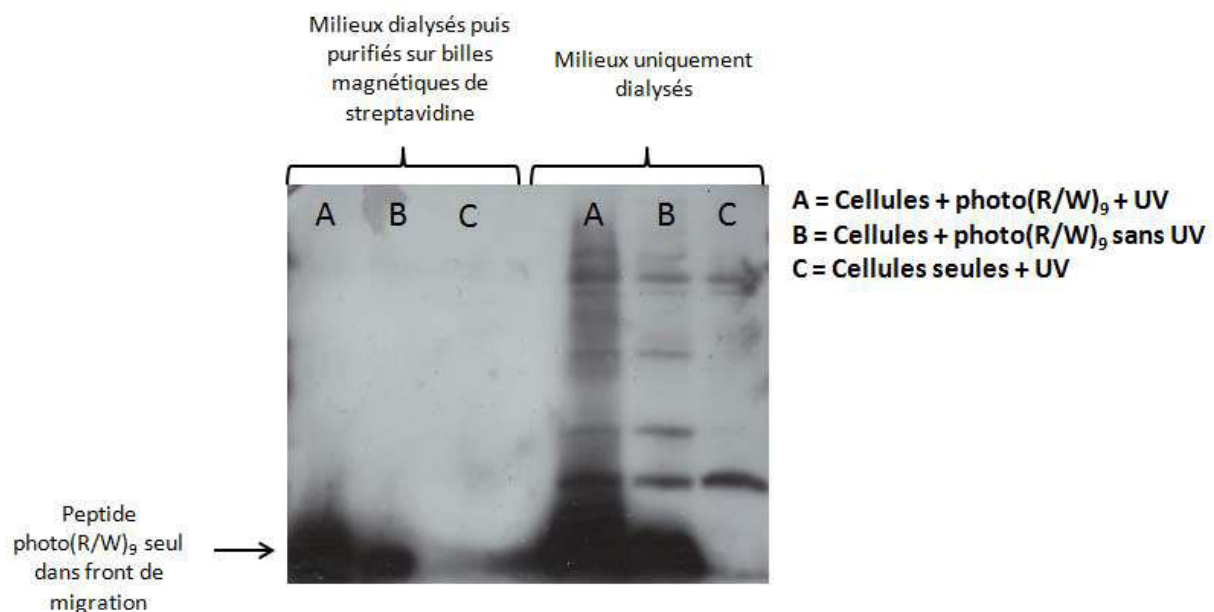
### Conclusion :

Ainsi cette première expérience complémentaire nous donne accès aux informations suivantes :

- Le peptide photo(R/W)<sub>9</sub>, vraisemblablement du fait de son caractère amphiphile et polycationique semble interagir avec un **nombre important de partenaires d'interactions**. Il est difficile d'imaginer que toutes ces interactions correspondent à un rôle biologique, ainsi beaucoup seraient vraisemblablement liées à des interactions non-spécifiques. Ces interactions peuvent à la fois correspondre à des liaisons hydrophobes avec les trois résidus tryptophane de photo(R/W)<sub>9</sub> et à des interactions électrostatiques connues pour être des liaisons plus fortes ( et de plus longue portée) du fait de la présence de six résidus arginines positivement chargés.
- Les **interactions non-covalentes** entre le peptide photo(R/W)<sub>9</sub> et un grand nombre de protéines sont suffisamment fortes pour être **maintenues au cours de la migration sur gel dénaturant 1D SDS-PAGE** (1% de SDS à la fois dans le gel et dans le tampon de migration).
- La **part des espèces photocross-linkées est faible** parmi l'ensemble des complexes photo(R/W)<sub>9</sub>-protéines (associations covalentes et non-covalentes). Cette contribution semble être un peu plus importante dans la partie supérieure de la piste A qui est plus sombre sur le western blot que pour la piste B.

### ➤ Est-il possible de purifier ces complexes photocross-linkés entiers sur billes ?

Pour essayer de répondre à cette question nous avons utilisé les mêmes milieux réactionnels A, B et C que nous avons cette fois dialysés, afin d'éliminer un maximum de peptide photo(R/W)<sub>9</sub> non covalamment lié (cut-off de 5 kDa), et purifiés sur billes magnétiques de streptavidine avant de les déposer sur gels et de les transférer sur membrane de nitrocellulose pour la réalisation de western blot (cf. Figure 59).



**Figure 59 : Image d'un western blot réalisé avec de la streptavidine HRP.** Les trois pistes de gauche correspondent aux milieux réactionnels A, B et C dialysés et soumis à une purification sur billes magnétiques de streptavidine. Les trois pistes de droite correspondent à ces mêmes échantillons mais non purifiés sur billes.

### Résultats :

Nous constatons immédiatement que les pistes de gauche, correspondant aux milieux purifiés sur billes, ne présentent aucune bande à l'exception de celle, en bas du gel, correspondant au peptide (R/W)<sub>9</sub> seul. Il n'y a donc soit pas d'espèces photocross-linkées, soit pas d'enrichissement de celles-ci sur les billes. Cependant, la révélation en western blot du peptide photo(R/W)<sub>9</sub> au niveau du front de migration montre que la streptavidine-HRP est capable d'interagir avec la biotine du peptide seul dans cette expérience. Si on s'intéresse à présent aux pistes situées à droite, nous observons que la piste A est globalement plus foncée que la piste B. La dialyse permettrait donc d'éliminer une partie du peptide photo(R/W)<sub>9</sub> seul et de limiter la quantité de complexes non-covalents peptide photo(R/W)<sub>9</sub>-protéines visibles pour la piste B et d'ainsi mieux mettre en évidence l'existence d'espèces photocross-linkées dans la piste A. Cependant certaines bandes demeurent intenses pour la piste B. Correspondraient-elles à des partenaires d'interaction de forte affinité de photo(R/W)<sub>9</sub> ? Parmi ces bandes au moins deux sont visibles dans la piste C et pourraient donc plutôt correspondre à un manque de spécificité de la streptavidine-HRP utilisée.

### Conclusion :

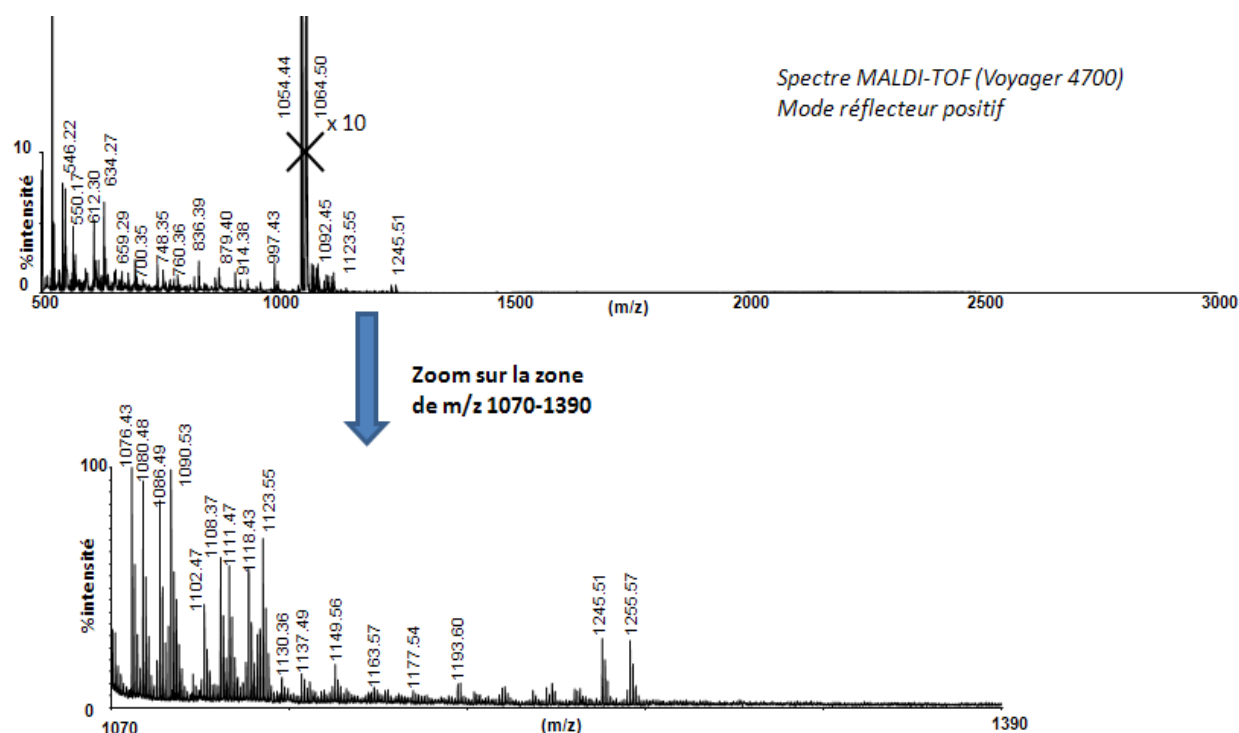
- Une quantité non négligeable d'espèces photocross-linkées semble pouvoir être formée lors de l'irradiation des cellules en présence du peptide photo(R/W)<sub>9</sub>.
- La dialyse permet d'éliminer une partie du peptide photo(R/W)<sub>9</sub> non covalamment lié à des protéines.
- La purification d'affinité des espèces photocross-linkées entières est inefficace.

### ➤ *L'identification de partenaires d'interactions sera-t-elle possible avec la branche b à l'issue de l'optimisation des conditions de digestion ?*

Après différents essais et modifications du protocole, nous sommes parvenus à une digestion complète du peptide (R/W)<sub>9</sub> et donc *a priori* à des conditions pour lesquelles la digestion des protéines du lysat cellulaire est également correcte. Les modifications essentielles du protocole de départ sont un **changement de la composition du tampon de lyse** qui présente une concentration en SDS beaucoup plus faible et également **l'utilisation d'un volume de tampon de lyse considérablement réduit** (Les protocoles de départ et ceux adaptés à l'issue de ces optimisations sont présentés dans la partie matériel et méthodes III. 7).

A l'issue de la digestion trypsique, le milieu est purifié sur billes magnétiques de streptavidine puis les espèces retenues sur ces billes sont éluées et analysées par spectrométrie de masse MALDI-TOF (Figure 60).

## Résultats : Vers la capture de partenaires par photocross-linking in cellulo



**Figure 60 : Spectre MALDI-TOF d'un digest de lysat cellulaire (cellules photoirradiées en présence de photo(R/W)<sub>9</sub>) après purification sur billes de streptavidine. Matrice utilisée CHCA.**

### Résultats :

- A l'issue de la digestion et de la purification par affinité du milieu de photocross-linking *in cellulo*, l'espèce observée dans le spectre MALDI-TOF de manière extrêmement majoritaire est le peptide photo(R/W)<sub>9</sub> après digestion complète à la trypsine (motif Biot-G<sub>5</sub>(<sup>1</sup>H/<sup>2</sup>H)-K(pBz)-R).
- Quelques doublets de m/z supérieurs sont observables dans la gamme 1100-1300 et sont susceptibles de correspondre à des espèces photocross-linkées. Cependant leur très faible intensité n'a pas permis d'envisager leur analyse avec le système NanoLC-ESI-MS/MS ni d'obtenir d'informations par fragmentation de ces doublets en MALDI-TOF/TOF.

De plus, on peut anticiper d'après la masse observée, que ces ions correspondant au motif Biot-G<sub>5</sub>(<sup>1</sup>H/<sup>2</sup>H)-K(pBz)-R lié covalamment avec, au maximum, deux à trois résidus ne pourront pas conduire à l'identification de partenaires d'interaction.

### Conclusion :

Même à l'issue des optimisations de digestion, il apparaît que la branche b de la stratégie initiale reposant sur une approche bottom-up ne peut pas permettre l'identification de partenaires d'interactions du peptide photo(R/W)<sub>9</sub>.

Nous savons avec les expériences réalisées *in vitro* que la réaction de photocross-linking n'est pas en cause (*cf.* Figure 42). Cependant il apparaît que la **diversité des partenaires d'interactions** suggérée par les expériences réalisées avec la branche a. conduisent à une **dilution importante du signal MS**.

## Résultats : Vers la capture de partenaires par photocross-linking *in cellulo*

Pour pouvoir espérer identifier un partenaire du peptide photo(R/W)<sub>9</sub>, il faudrait que ce dernier ait une particulièrement bonne affinité avec le peptide et qu'il soit le seul ou un des seuls partenaires intracellulaires du peptide. Pour limiter la dilution du signal MS des peptides photocross-linkés, il faut être en mesure de décomplexifier le mélange. C'est ce que nous avons cherché à faire par la suite.

### c. Adaptation de la stratégie initiale : vers un fractionnement cellulaire.

Les différentes expériences de photocross-linking *in cellulo* réalisées ont mis en évidence la **nécessité d'une séparation préalable** (sur gel 1D SDS-PAGE par exemple) à la digestion des espèces photocross-linkées de manière à avoir (1) une information sur la masse de la protéine partenaire et surtout de façon à (2) limiter la dilution du signal en spectrométrie de masse du fait de la présence d'une trop grande diversité de peptides photocross-linkés.

Ces expériences ont également mis en évidence l'impossibilité de purifier par affinité les complexes photocross-linkés entiers (malgré la grande spécificité et affinité de la streptavidine pour la biotine, cette dernière peut être masquée par les partenaires associés au peptide photo(R/W)<sub>9</sub> biotinylé).

Par conséquent la stratégie initiale s'avère être non adaptée à la recherche de partenaires du peptide photo(R/W)<sub>9</sub>.

A partir de ces informations nous avons réfléchi à une stratégie alternative combinant les étapes des deux branches de la stratégie initiale et intégrant les informations sur la localisation d'espèces biotinylées pouvant être obtenues en western blot. Nous avons également décidé de travailler ici, non pas avec des cellules entières mais avec seulement une partie du matériel cellulaire obtenu après une étape de **préfractionnement par ultracentrifugation différentielle** (fractions cellulaires : cytosol, lysat ou membranes). Le protocole de préfractionnement cellulaire est donné dans la partie matériel et méthodes III.7. Cette nouvelle stratégie a été schématisée Figure 61.

Dans cette stratégie l'utilisation de la technique de western blot en parallèle de la réalisation d'un gel 1D SDS-PAGE renseigne sur la contribution du photocross-linking en permettant de visualiser la présence d'espèces photocross-linkées.



## Résultats : Vers la capture de partenaires par photocross-linking *in cellulo*

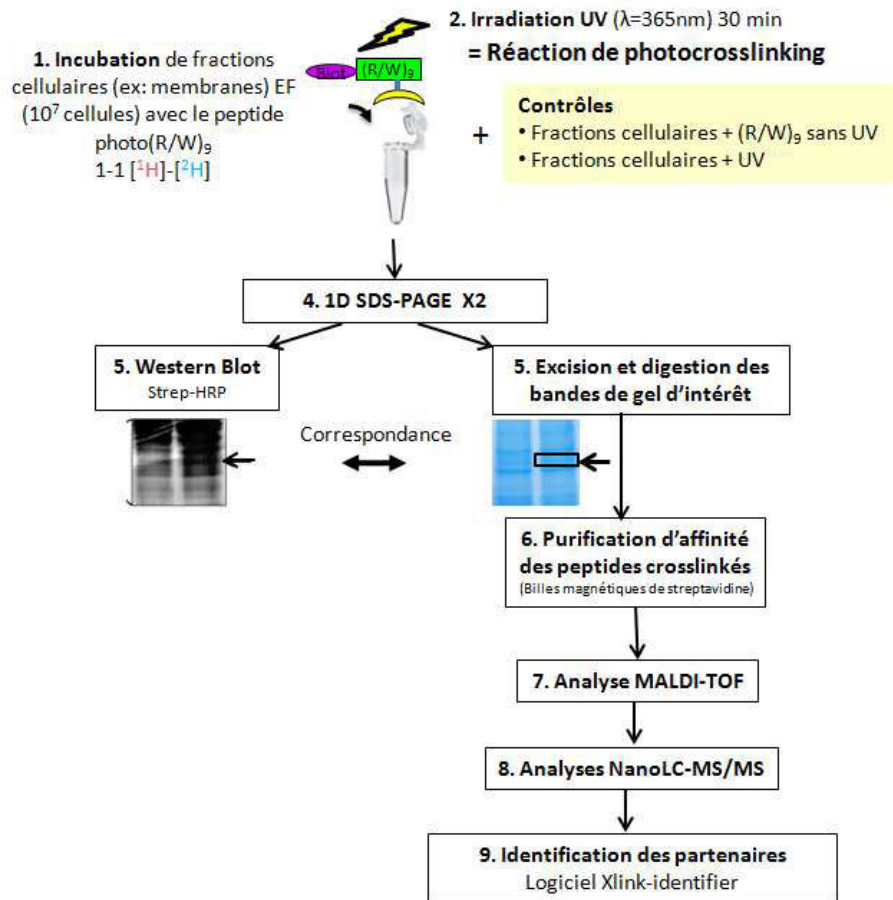


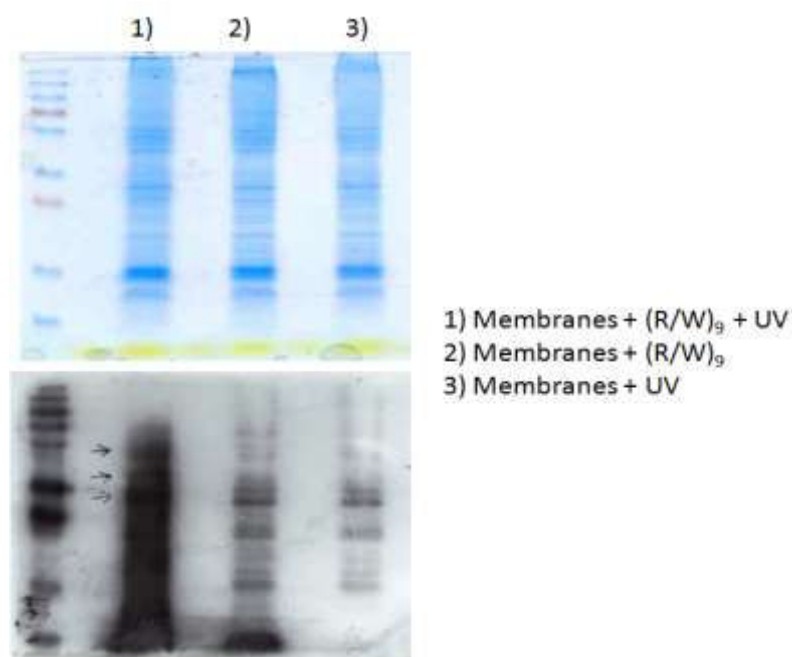
Figure 61 : Schéma de la stratégie "adaptée" (N°2) de photocross-linking *in cellulo*.

### Exemple avec la fraction membranaire

J'ai choisi de vous présenter ici uniquement l'application de cette stratégie N°2 au cas de la fraction membranaire (le protocole détaillé de l'expérience est donné dans la partie matériel et méthodes III.7).

A l'issue de l'incubation et de l'irradiation UV, les échantillons sont migrés sur 2 gels SDS-PAGE, l'un pour une coloration au bleu de Coomassie et l'autre pour un électrotransfert sur membrane de nitrocellulose avant western blot avec la streptavidine-HRP.

## Résultats : Vers la capture de partenaires par photocross-linking in cellulo



**Figure 62 : Gel 1D SDS-PAGE coloré au bleu de Coomassie et Western blot avec de la streptavidine HRP pour l'expérience de photocross-linking sur membranes.**

Nous constatons (Figure 62) que la piste correspondant à l'échantillon de membranes photoirradié en présence de peptide est très sombre en comparaison avec les pistes 2 et 3 indiquant la présence d'un grand nombre d'espèces biotinylées, vraisemblablement liées de façon covalente avec le peptide photo(R/W)<sub>9</sub>.

Nous avons excisé quelques bandes, à la même hauteur pour chacune des pistes, que nous avons ensuite digérées et purifiées par affinité pour voir si nous pouvions mettre en évidence des peptides photocross-linkés. Cependant nous ne sommes pas parvenus à observer de doublets caractéristiques de la présence d'espèces photo-crosslinkées.

### Conclusion et perspectives :

Les expériences de photocross-linking menées sur cellules entières et sur fractions cellulaires ne nous ont pas permis d'identifier de partenaires intracellulaires ou membranaires du peptide photo(R/W)<sub>9</sub> et ne nous ont donc pas aidé à comprendre son mode action sur les cellules tumorales EF.

Malgré tout, ces expériences nous ont apporté des informations : (1) Le peptide photo(R/W)<sub>9</sub> semble avoir un grand nombre de partenaires dans la cellule ; (2) il est capable de s'associer non covalamment en interagissant fortement, et vraisemblablement *via* ses charges positives, avec différentes protéines intracellulaires ou membranaires.

Ces expériences nous ont également amenés à apporter des améliorations à la stratégie de photocross-linking initiale et nous pouvons imaginer que cette stratégie pourrait être utilisée ultérieurement pour la recherche de partenaires de biomolécules impliquées dans un nombre plus limité d'interactions. En effet, le peptide photo(R/W)<sub>9</sub> a des propriétés physicochimiques particulières qui font que ce peptide a une forte propension à l'auto-assemblage mais aussi à faire des liaisons électrostatiques et hydrophobes.



## Chapitre 6 : Pull down et photocross-linking

Lors des expériences de photocross-linking *in cellulo* décrites dans le chapitre 5, nous avons pu faire plusieurs constatations : (1) il est impossible, dans nos conditions, de capturer sur billes de streptavidine les complexes covalents entiers, biotinylés sur la partie N-terminale du CPP ; (2) le CPP photo(R/W)<sub>9</sub> est capable de rester associé de manière non-covalente avec de nombreuses protéines des lysats de cellules EF malgré leur migration effectuée sur gel d'électrophorèse 1D SDS-PAGE en conditions dénaturantes, indiquant une forte affinité pour ces protéines.

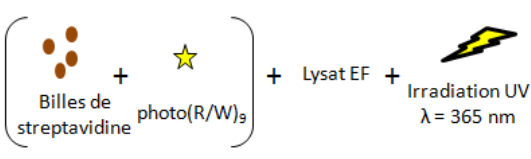

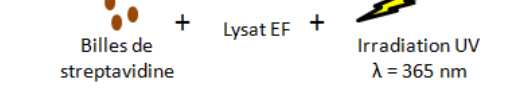
Ces éléments nous ont conduits à réaliser des expériences de pull-down. En effet, puisque nous ne pouvons pas récupérer les complexes photocross-linkés entiers sur billes après leur formation, nous avons choisi de réaliser la capture des partenaires directement sur les billes fonctionnalisées avec le peptide photo(R/W)<sub>9</sub> qui sert alors d'appât .

### I. Présentation de l'expérience de pull-down associée au photocross-linking

La technique de pull-down fait partie des techniques de purification d'affinité que nous avons présentées dans le chapitre 1. Brièvement, elle consiste à immobiliser une biomolécule d'intérêt sur un support et à incubier l'ensemble avec un mélange complexe de partenaires potentiels. Ici, la molécule d'intérêt est le peptide photo(R/W)<sub>9</sub>. Il est immobilisé sur des billes magnétiques de streptavidine *via* son étiquette biotine et incubé avec un lysat cellulaire EF. Tandis que les interactions non-covalentes stables pourront être préservées lors des lavages des billes, les interactions faibles risquent d'être éliminées lors de cette étape. Pour chercher à éviter la perte de ces interactions transitoires qui peuvent jouer un rôle crucial d'un point de vue biologique, nous avons ajouté une étape de photoirradiation (*cf.* chapitre 3) au protocole classique de pull-down. Cette étape de photoirradiation intervient à l'issue de l'incubation avec le lysat cellulaire EF et avant le lavage des billes. Elle est rendue possible par la présence de la sonde benzophénone présente dans la séquence du peptide photo(R/W)<sub>9</sub>.

Pour chaque expérience nous avons ainsi travaillé avec trois conditions résumées sur le Tableau 10. Le protocole détaillé de l'expérience est donné dans la partie matériel et méthodes III. 8.

**Tableau 10 : Présentation des différents échantillons de l'expérience de pull-down associée au photocross-linking**

Milieu 1	Pull-down+ photocross-linking	
Milieu 2	Pull-down	
Milieu 3	Contrôle négatif	

Ces trois échantillons contiennent la même quantité de billes magnétiques de streptavidine qui sont incubées avec des volumes identiques de lysats EF soit une quantité de protéines de l'ordre de 1 mg. Pour les milieux 1 et 2, les billes sont préalablement fonctionnalisées avec peptide photo(R/W)<sub>9</sub>. Les milieux 1 et 3 sont photoirradiés à l'issue de l'incubation avec le lysat EF.

Comme nous l'avons vu dans le chapitre 2, partie I.4.c, il est nécessaire de prendre un certain nombre de précautions pour réaliser ce type d'expérience. En particulier, il faut choisir des conditions de lyse permettant de conserver la structure tridimensionnelle des protéines ainsi que les complexes protéiques. Nous avons ainsi veillé à choisir un tampon non-dénaturant à base de détergent non-ionique, le Nonidet P-40 (ou NP-40), pour la préparation des lysats de cellules EF. Nous avons ensuite fait en sorte de ne pas faire subir de changements de température importants aux lysats cellulaires en les maintenant constamment à 4 °C (à l'exception de l'étape de photoirradiation réalisée à température ambiante).

Nous avons constaté que le montage de photoirradiation présenté chapitre 5 n'était pas adapté à notre expérience de pull down + photocross-linking. En effet, nous avons observé que les billes magnétiques de streptavidine ont tendance à s'agréger lorsqu'elles subissent une photoirradiation, quelle que soit la durée d'irradiation et bien qu'étant placées dans un bain eau/glace. Ce problème d'agrégation, certainement dû à une augmentation de température locale même faible ou, au contraire, à une température trop faible dans le bain eau/glace, a pu être contourné en utilisant un système Rayonet mis à disposition par l'Institut Parisien de Chimie Moléculaire (UMR 7201, Dr. Cyril Olivier). Ce montage commercialisé par Southern New England company est équipé de 16 néons disposés de façon circulaire et émettant autour de la longueur d'onde  $\lambda = 350$  nm. Une ventilation empêche l'augmentation de la température dans le montage et permet ainsi d'éviter de placer les échantillons à photoirradier dans la glace.

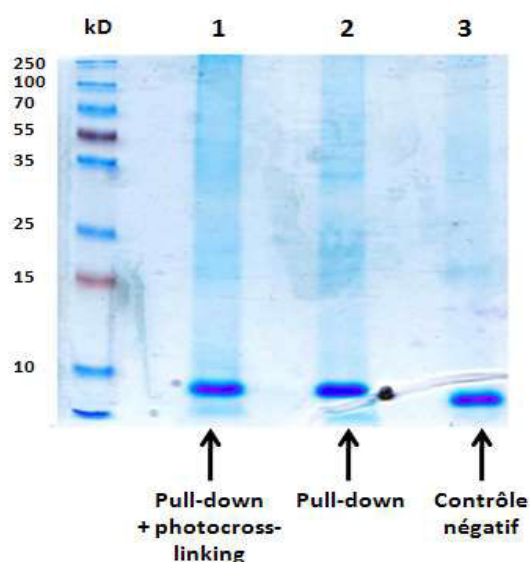
Après incubation avec le lysat EF, suivie ou non d'une étape de photoirradiation, les billes sont lavées avec un tampon doux de manière à éliminer une part importante des interactions non spécifiques. Les billes sont ensuite éluées et les éluats déposés sur trois pistes d'un gel d'électrophorèse 1D SDS-PAGE. Les trois pistes sont ensuite systématiquement découpées, digérées par la trypsine et

analysées en spectrométrie de masse haut-débit. Les échantillons sont donc soumis à une stratégie dite « bottom-up ».

## II. Résultats obtenus

### 1. Electrophorèse sur gel 1D-SDS-PAGE

L'expérience a été répétée entièrement deux fois. Ci-dessous (Figure 63) se trouve l'exemple d'un gel coloré au bleu de Coomassie, obtenu pour une de ces expériences.



**Figure 63 : Gel d'électrophorèse 1D SDS-PAGE coloré au bleu de Coomassie.** La piste de gauche correspond aux marqueurs de masses moléculaires. Les masses moléculaires étant indiquées sur la gauche, en kilodaltons (kDa). Sur les pistes numérotées 1, 2 et 3 ont été déposées et migrées les protéines éluées des billes magnétiques de streptavidine correspondant respectivement au milieu de pull-down + photocross-linking (1), au pull-down simple (2) et au contrôle négatif (3). Les bandes intenses présentes dans la partie inférieure du gel correspondent au monomère de la streptavidine décroché des billes (MM ~ 12 kDa).

A première vue, nous constatons que la coloration de la piste 1 est légèrement plus intense que celle de la piste 2, elle-même plus colorée que la piste 3. Il semble donc que le peptide (R/W)<sub>9</sub> ait permis de capturer un certain nombre de protéines (piste 2) et que la photoirradiation permettrait de maintenir un nombre plus important de protéines à l'issue de l'étape de lavage des billes. Ces premières observations qualitatives, confirmées par une coloration à l'argent (pas montré), vont donc dans le sens attendu.

### 2. Résultats des analyses de spectrométrie de masse haut débit.

Le traitement des données brutes obtenues pour l'analyse des différentes pistes de gel après digestion à la trypsine conduit à l'identification des protéines retenues sur les billes à l'issue des lavages pour les trois conditions. Les nombres de protéines et de peptides identifiés étant proches pour les deux réplicats biologiques, nous avons choisi de présenter les résultats moyennés sur ces deux expériences (Tableau 11).

**Tableau 11 : Nombres moyens de peptides et de protéines identifiés pour les différentes pistes pour les deux réplicats biologiques.** Piste 1 : milieu de Pull-down+ Photocross-linking ; piste 2 : milieu de pull-down ; piste 3 : milieu contrôle négatif. FDR : False Discovery Rate.

Piste	Nombre moyen de protéines identifiées	Nombre moyen de peptides identifiés (1%FDR)
1	258	670
2	167	389
3	117	280

Les résultats obtenus en termes de nombre de protéines et de peptides identifiés vont dans le même sens que les observations faites sur le gel d'électrophorèse 1D SDS-PAGE : il y a davantage de protéines piégées sur les billes en présence du peptide photo(R/W)<sub>9</sub> (pistes 1 et 2), et ce nombre augmente à nouveau de manière très significative lorsqu'une étape de photoirradiation est réalisée (piste 1).

Ces résultats nous indiquent également qu'un nombre conséquent de protéines interagissent de manière non-spécifique avec les billes magnétiques (polymères de revêtement et streptavidine). En effet, le nombre de protéines identifiées dans le contrôle négatif est supérieur à 100.

Ce tableau récapitulatif (Tableau 11) présentant des données moyennées est donné à titre indicatif. En effet, pour l'évaluation des partenaires protéiques potentiels, nous avons gardé séparées les listes de protéines obtenues pour les deux expériences.

### **III. Interprétation des données : Quelles sont les « vraies » interactions?**

#### **1. Génération des listes de partenaires d'interaction potentiels**

##### **a. Approche binaire : Présence ou absence d'une protéine dans une liste**

Pour déterminer les partenaires protéiques susceptibles d'interagir de manière spécifique avec le peptide photo(R/W)<sub>9</sub>, la première méthode envisageable est de soustraire des listes de protéines obtenues pour les milieux 1 et 2, les protéines identifiées dans la piste 3 contrôle correspondant aux interactions non-spécifiques des protéines des lysats avec les billes. La liste de protéines discriminantes obtenue constitue, dans ce cas, la liste des partenaires potentiels. Les protéines qui ne seraient vues que pour la piste 1, correspondraient à des protéines interagissant de manière transitoire ou faible avec photo(R/W)<sub>9</sub>.

Cependant cette méthode, qui peut être qualifiée d'approche binaire, nous est apparue trop approximative dans la mesure où elle ne tient pas compte de la façon dont sont identifiées ces



différentes protéines, c'est-à-dire du nombre de peptides identifiés et du nombre de spectres ayant conduit à l'identification de ces peptides ou de l'intensité des ions correspondants. Ainsi, par exemple, elle peut nous conduire à éliminer une protéine qui ne serait vue qu'avec un unique peptide dans le contrôle négatif alors qu'elle pourrait être identifiée avec un nombre plus important de peptides (ex. 30 peptides) dans un milieu de pull-down et par conséquent, correspondre à un partenaire d'intérêt.

### b. Approche de quantification par comptage spectral

Dans la mesure où nous ne souhaitons pas trier les listes de protéines simplement en termes de présence ou d'absence de celles-ci dans les différentes listes, nous nous sommes intéressés aux outils qui pourraient permettre une analyse plus fine de ces résultats. Dans le chapitre 2, partie I.4 c., nous avons vu combien il pouvait être difficile de distinguer les interactions dites spécifiques, c'est-à-dire ayant une fonction biologique, d'interactions non-spécifiques et nous avons présenté quelques-uns des outils développés pour faciliter cette tâche. L'outil qui nous semble actuellement le plus abouti est le logiciel SAINT pour *Significance Analysis of Interactome* développé conjointement par les équipes de Gingras A.C et de Nesvizhskii A. [135]. De plus, nous avons pu soumettre les listes de protéines obtenues facilement *via* l'interface web du logiciel sur le site CRAPome.org [288].

Le logiciel SAINT a été initialement développé pour **convertir des données de quantification en label free**, à savoir le nombre de spectres MS/MS  $X_{ij}$  pour une protéine proie  $i$  identifiée dans une expérience de purification d'affinité avec l'appât  $j$ , **en probabilité de vraie interaction**  $P(\text{vraie}/X_{ij})$  entre ces deux protéines. Les nombres de spectres MS/MS acquis pour chaque paire proie-appât sont modélisés selon une loi de probabilité mixte constituée de deux composantes : une loi de probabilité pour le nombre de fois où la paire correspond à une fausse interaction et une loi de probabilité pour le nombre de fois où la paire correspond à une vraie interaction. Le logiciel SAINT normalise le nombre de spectres MS/MS par rapport à la longueur des protéines et en fonction du nombre total de spectres acquis pour l'expérience de purification par affinité et considérés comme correspondant à des vraies interactions. Lorsque des contrôles négatifs sont disponibles, SAINT prend en compte la présence de protéines dans ce contrôle pour pénaliser la probabilité de vraie interaction de ces protéines avec l'appât. On parle ainsi d'approche semi-supervisée. Une présentation plus détaillée du modèle statistique utilisé est donnée dans la publication de Choi H. *et al.* [289].

A chaque fois il s'agit de comparer une condition « Expérience » avec une condition « Contrôle ». Nous avons réalisé deux recherches de ce type avec le logiciel SAINT :

- A : La comparaison des listes de protéines obtenues pour la condition Pull-down + photocross-linking avec les listes de la condition contrôle négatif (1 *versus* 3).
- B : La comparaison des listes de protéines obtenues pour la condition Pull-down avec les listes de la condition contrôle négatif (2 *versus* 3).

Les résultats des recherches sont présentés sous forme de tableaux donnant la liste des protéines de la condition « Expérience » avec les probabilités de vraie interaction correspondantes, calculées en comparant le nombre de spectres MS/MS acquis pour une protéine dans la condition « Expérience » et dans la condition « Contrôle ». A chaque fois, nous avons utilisé les résultats des deux réplicats

biologiques en parallèle et avons ainsi pu obtenir une **probabilité moyenne de vraie interaction notée AvgP**. Nous nous sommes basés sur cette valeur d'AvgP pour filtrer les données. Cependant le choix du seuil pour la valeur d'AvgP est tout à fait arbitraire. Le fait de choisir un seuil relativement bas permet de ne pas passer à côté de partenaires potentiellement intéressants. Il faut impérativement garder à l'esprit que ce logiciel vise à trier les protéines identifiées pour faciliter l'interprétation des données et que les protéines jugées intéressantes devront par la suite être soumises à des expériences complémentaires de biologie. Le logiciel fournit par conséquent des pistes possibles mais pas de preuves.

Les listes de protéines partenaires potentielles présentant une AvgP supérieure à 0,75 pour les différentes comparaisons sont présentées en annexes N° 3 et 4 car de longueurs relativement importantes.

## 2. Interprétation avec les outils d'enrichissement de termes de *gene ontology*

Pour l'interprétation de ces listes de protéines avec une AvgP > 0,75, nous avons essayé de faire appel aux logiciels permettant de réaliser l'enrichissement de termes de *Gene Ontology (GO)* (cf. Chapitre 2, partie III, 2) tels que STRING [195], [196] et DAVID [193], [194]. L'enrichissement d'un terme de *gene ontology* est traduit par un paramètre statistique appelé p-value : plus cette valeur est petite plus l'enrichissement est significatif pour ce terme dans la liste soumise. Le processus biologique le plus enrichi s'est avéré être la traduction des protéines (p-value =  $1,7 \cdot 10^{-11}$  pour A ; p-value =  $2,6 \cdot 10^{-8}$  pour B). En terme de fonction moléculaire c'est la « liaison à l'ARN poly(A) » qui est arrivée en tête pour les deux comparaisons (p-value =  $2,6 \cdot 10^{-18}$  pour A ; p-value =  $5,4 \cdot 10^{-20}$  pour B). Cependant nous nous sommes rendus compte que nous ne pouvions pas donner de crédit à ces résultats d'enrichissement car ces termes de GO « traduction » et « liaison à l'ARN poly(A) » arrivaient également parmi les premiers pour la liste de protéines du contrôle négatif seul.

Les protéines responsables de l'enrichissement de ces termes de GO sont majoritairement les protéines ribosomales. Or, ces protéines ribosomales sont souvent décrites comme des contaminants des expériences de purification d'affinité [115]. Nous n'avons pas pu nous affranchir complètement de ces protéines malgré l'utilisation d'une condition contrôle négatif avec les billes seules.

Il peut y avoir plusieurs raisons à cela :

- Le **nombre de réplicats biologiques n'est pas assez élevé** pour couvrir toutes ces protéines contaminantes dans les contrôles négatifs ;
- Le peptide (R/W)<sub>9</sub> interagit avec ces protéines liant l'ARN même si ces **interactions ne sont pas spécifiques** dans le sens où elles n'induisent pas l'effet biologique donné. Ces interactions peuvent être liées au fait que le peptide (R/W)<sub>9</sub> avec son alternance de résidus arginine et tryptophane est un concentré de « points chauds » (contribution de plus de 2 kcal/mol dans la formation d'une interaction). En effet, Bogan A. et Thorn K. qui se sont intéressés à l'anatomie des points chauds dans les interfaces d'interactions protéine-protéine ont montré que seulement trois acides aminés apparaissaient avec une fréquence de plus de 10 % dans les points chauds : 21 % pour le tryptophane, 13,3 % pour l'arginine et

12,3 % pour la tyrosine [290]. De plus, l'ARN étant chargé négativement, il va pouvoir former de nombreuses interactions électrostatiques avec le peptide (R/W)<sub>9</sub>

- Finalement, l'interaction du peptide (R/W)<sub>9</sub> avec ces protéines liant l'ARN pourrait avoir un **rôle biologique réel**. Cet argument s'appuie notamment sur un article de Bayer T. *et al.* qui montre que les protéines avec des motifs riches en arginine, tel que le peptide Tat, lient spécifiquement l'ARN et sont impliquées dans différents processus de régulation de l'ARN [291].

Dans ce contexte, il n'a donc pas été possible de tirer d'informations pertinentes des enrichissements de termes de GO.

Nous nous en sommes donc remis à un examen manuel des listes des protéines obtenues. Il est important de rappeler à ce stade qu'un des défis majeurs de la discipline est l'interprétation des listes de protéines générées, la capacité de donner un sens biologique aux données de protéomique. Cette tâche est d'autant plus ardue que notre biomolécule d'intérêt, le peptide vecteur (R/W)<sub>9</sub> est une molécule synthétique de conception rationnelle et qu'elle n'a donc *a priori* aucun partenaire « naturel » dans la cellule.

### 3. Examen manuel des listes de partenaires d'interaction potentiels

L'examen de ces listes nous a livré au moins un élément particulièrement intéressant : En effet, parmi les protéines identifiées avec un très bon score de probabilité de vraie interaction pour la comparaison des expériences de pull-down avec les contrôles négatifs (B = 2 vs 3), se trouve la protéine *Ewing Sarcoma breakpoint region 1* ou EWS avec un AvgP = 0,99 (*cf.* chapitre 1 partie III.2).

Dans chacun des deux réplicats biologiques, la protéine EWS a été identifiée avec 2 peptides uniques eux-mêmes identifiés grâce à deux spectres MS/MS chacun soit un total de 4 spectres par réplicats contre aucun dans les deux contrôles négatifs (*cf.* Tableau 12). Une autre protéine a également particulièrement retenu notre attention : le polypeptide 1 DEAD box qui est un motif très conservé dans les protéines ARN hélicases.

**Remarque** : Il convient de rappeler que la protéine EWS est toujours exprimée dans les cellules tumorales EF. En effet, les cellules EF résultent de la transduction de fibroblastes 3T3 par un rétrovirus contenant un gène de fusion codant pour la protéine EWS-FLI1 ainsi qu'un gène de résistance à l'antibiotique puromycine. Ainsi, lorsque les fibroblastes 3T3 ont intégré le nouveau gène au niveau d'un de leurs deux allèles ils vont subir la transformation oncogénique de EWS-FLI1 (devenir des cellules appelées EF) et acquérir la résistance à la puromycine qui leur permettra d'être sélectionnés. Néanmoins le second allèle continuera à coder pour la protéine EWS native.

**Tableau 12 : Extrait du tableau de résultats obtenu avec le logiciel SAINT pour l'expérience de Pull-down simple.** colonne 1 : Numéro de la protéine, sachant que les protéines ont été classées dans l'ordre décroissant de leurs valeurs de probabilité moyenne de vraie interaction (AvgP) ; colonne 2 : Numéro d'accès Uniprot ; colonne 3 : nom communément donné à la protéine ; colonne 4 : référence des expériences de Pull-down ; colonne 5 : nombre de spectres MS/MS ayant permis l'identification de la protéine dans chaque réplicat ; colonne 6 : nombre total de spectres MS/MS ayant permis l'identification de la protéine ; colonne 7 : nombre de réplicats de l'expérience ; colonne 8 : probabilité de vraie interaction calculée par le logiciel SAINT pour chaque expérience ; colonne 9 : nombre de spectres MS/MS ayant conduit à l'identification de la protéine dans les contrôles négatifs ; colonne 10 : probabilité moyenne de vraie interaction calculée sur les 2 réplicats. La ligne du tableau correspondant à la protéine partenaire d'intérêt, EWS, est surlignée en jaune ainsi que celle correspondant au polypeptide DEAD box qui a également retenu notre attention.

	1	2	3	4	5	6	7	8	9
N°	Bait	PreyGene	IP	Spec	SpecSum	NumRep	ctrlCounts	AvgP	MaxP
...									
28	Pull-down	lamin A	Exp2 Exp1	4 5	9	2	0 0	1	1
29	Pull-down	ribosomal protein L10	Exp2 Exp1	8 5	13	2	0 0	0,9987	1
30	Pull-down	heat shock protein 8	Exp2 Exp1	5 6	11	2	0 0	0,9975	1
31	Pull-down	similar to ribosomal protein L13	Exp2 Exp1	3 4	7	2	0 0	0,9975	1
32	Pull-down	cytoskeleton-associated protein 4	Exp2 Exp1	5 25	30	2	0 0	0,995	0,998
33	Pull-down	ATP synthase, H+ transporting mitochondrial F1 complex, beta subunit	Exp2 Exp1	23 5	28	2	0 0	0,9925	1
34	Pull-down	Ewing sarcoma breakpoint region 1	Exp2 Exp1	4 4	8	2	0 0	0,99	0,993
35	Pull-down	DEAD (Asp-Glu-Ala-Asp) box polypeptide 1	Exp2 Exp1	2 5	7	2	0 0	0,9737	0,993
36	Pull-down	similar to 60S ribosomal protein L3 (J1 protein)	Exp2 Exp1	2 8	10	2	0 0	0,9688	0,995
37	Pull-down	40S ribosomal protein S3a	Exp2 Exp1	6 2	8	2	0 0	0,9688	0,988
38	Pull-down	ribosomal protein S3	Exp2 Exp1	3 14	17	2	0 0	0,9675	0,998
39	Pull-down	vimentin	Exp2 Exp1	2 3	5	2	0 0	0,965	0,983
...									

En dehors de ces expériences de pull-down avec le peptide photo(R/W)<sub>9</sub>, nous ne sommes jamais parvenus à identifier les protéines EWS, EWS-Fli1 et Fli1. C'est le cas notamment pour les expériences de protéomique globale réalisées sur les cellules EF et qui seront abordées dans le chapitre suivant. Ces protéines ne faisaient jamais partie des près de 1700 protéines identifiées et quantifiées pour cette lignée cellulaire EF. Ainsi il semblerait que la **protéine EWS soit très faiblement abondante dans les lysats cellulaires EF ce qui rend son identification dans les expériences de pull-down d'autant plus intéressante**. Pour être en mesure d'identifier ces protéines faiblement exprimées, un fractionnement plus poussé ou le recours à un mode d'acquisition indépendant des données (DIA) est envisageable.

## IV. Signification biologique des résultats obtenus

### 1. Une interaction avec la protéine EWS

La protéine EWS a été identifiée dans les expériences de pull-down avec deux peptides correspondant aux résidus [269-304] et [373-386] (voir tableau d'attribution des spectres MS/MS en annexes N°5). Dans le chapitre 1, nous avons vu que le sarcome d'Ewing était initié par une translocation chromosomale impliquant le gène EWS du chromosome 22 et le gène FLI-1 du chromosome 11. Cette translocation chromosomale aboutit à l'expression d'une protéine de fusion

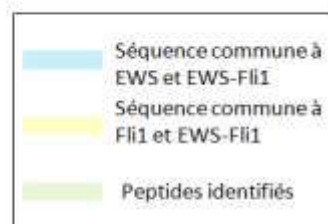
appelée EWS-FLI1 qui correspond à la partie N-terminale de la protéine EWS (résidus 1-265) et à la partie C-terminale de la protéine FLI1 (résidus 211-476) (Figure 64).

### EWS 656 acides aminés

```

1 mactdyatys qaaaqqyya ytaqptqgya qttqayqqq ytygqptdv sytqaqtat
61 yqqtayaty gqptqyttp tapqayqpv qygtgdyt tstatvtttq syaaqaygt
121 qpaypqqq paataptqp dqnkptetq pqaetggynq palgyqqany sypqvgayp
181 mqpvtappay pptyastqp taidqssayq qntygqssy gqqssyqqq syqqqptay
241 ppqtgssyqa pgyyqqqss yqqqssfrq hpsmgyvyy eaggfagpge nramsgpdr
301 ggrggfdrg gmarggrrg rggmgager ggnkpggpm degpdldlg pvdpededn
361 saiyvqglnd svtlldladf fkqcgvrkmn xrtggpmih yldketgkpk gdatvayedp
421 ptakaavewf dgkdfqgskl kvslarkkpp mnsargglpp regrgmfppl rggpgggpgp
481 ggmrgmrgg gdrgggfpq gprgargnpa gggnvqhrag dwqcpnpgc nqnfawrtec
541 nqckepkpeg flpppfpppg gdrgrggpg mrggrgglnd rggpggmfrg grggdrggrf
601 ggrgmrggf gggrrggpg ppgplneqng grrggrrggp kmkqgehrq rrdrgy

```



### FLI1 452 acides aminés

```

1 mdgtikeals vvaddqlfd saygaahlp kadmtasgp dygqphkinp lppqqewinq
61 pvrnvkrey dhmgasreap vdcavakcak lvgggsenpm nysymdekn qppppnmta
121 errvivpadp tiwtqehvrq wlewaikya imeidteffq nmdgkelcm nkedflrat
181 lynteவில்லish lsylrealll synttshtdq eerlvkedp sydsrvrgaw gnnmnglnk
241 spplggagti sknteqrqp dpyqilgptc erlanpggg iqlwqlllel lsdanasci
301 twegtngfak mtdpdevarr wgerkekpm nydklaralr yyydknmtk vhgkryaykf
361 dfhglagql phptessmyk ypedisympa yhaqqkvnf vpphpsmvp tassffgaa
421 qwtstpggi ypnnavprhp nthvpshlga yy

```

### EWS-FLI1 476 acides aminés

```

1 mactdyatys qaaaqqyya ytaqptqgya qttqayqqq ytygqptdv sytqaqtat
61 yqqtayaty gqptqyttp tapqayqpv qygtgdyt tstatvtttq syaaqaygt
121 qpaypqqq paataptqp dqnkptetq pqaetggynq palgyqqany sypqvgayp
181 mqpvtappay pptyastqp taidqssayq qntygqssy gqqssyqqq syqqqptay
241 ppqtgssyqa pgyyqqqss yqqqssplgg qttaknteq spqdpqyil qptsarlanp
301 gggqilwqf llelladnan aacitwegtn gefkutdpde varzwgqrks kphmnydkls
361 rairyydykn imtkvngkry aykdfdhgia qalqphptes smykypedis ympayhaqq
421 kvnfvpphs mpvtassff gaesqywtsp tgglyppnprv prhpnthvps hlqsy

```

265 premiers a.a de EWS

+

211 derniers a.a de FLI1

**Figure 64 : Présentation des séquences en acides aminés des protéines EWS, FLI1 et EWS-FLI1 et de leur degré d'homologie.**

Ainsi, les deux peptides identifiés font partie de la région de EWS qui n'est pas commune à la protéine EWS-FLI1 et cela nous permet de dire que c'est bien la protéine EWS qui a été piégée sur les billes greffées par le peptide photo(R/W)<sub>9</sub>. Cependant cela ne nous donne aucune indication sur l'interaction d'EWS avec les billes greffées par le peptide photo(R/W)<sub>9</sub>. En particulier, nous ne savons pas si la protéine EWS a été piégée *via* une interaction directe avec le peptide photo(R/W)<sub>9</sub> ou bien si c'est un partenaire d'interaction de EWS qui interagit directement avec photo(R/W)<sub>9</sub>. Dans l'hypothèse où le peptide photo(R/W)<sub>9</sub> interagit avec EWS directement, nous ne savons pas non plus si cette interaction implique la partie de séquence commune à EWS ou EWS-FLI1 ou bien la partie discriminante.

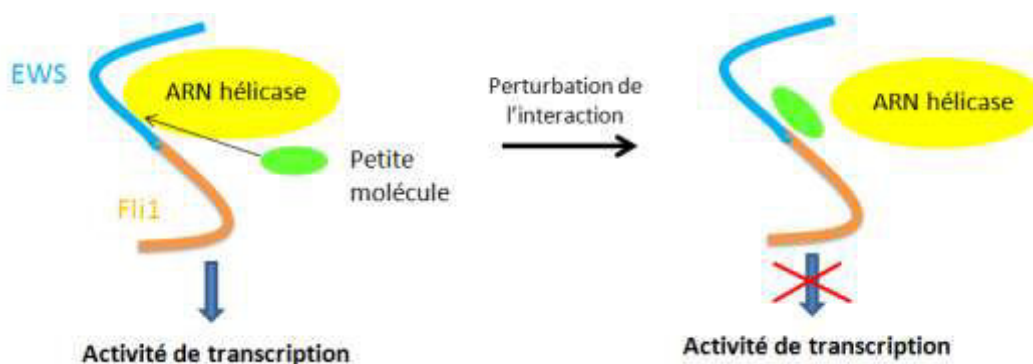
## 2. Possibilité d'une interaction avec EWS-FLI1

Comme nous l'avons indiqué précédemment, la protéine EWS, ainsi que les protéines FLI1 et EWS-FLI1 sont vraisemblablement des protéines faiblement exprimées ou, du moins, difficiles à identifier avec les moyens dont nous disposons.

La protéine EWS a été identifiée à deux reprises dans les expériences de pull down avec deux peptides dont aucun compris dans la séquence commune à EWS et EWS-FLI1. Cependant, on peut imaginer que si le peptide photo(R/W)<sub>9</sub> interagit avec EWS, il pourrait également interagir avec EWS-FLI1 étant donné l'homologie de séquence existant entre les deux protéines.

Il ne s'agit bien entendu que d'une supposition qui devra être testée mais qui mérite que l'on s'y attarde.

En effet, la protéine EWS-FLI1 joue un rôle absolument décisif dans le processus de tumorigénèse et constitue par conséquent une cible idéale pour le traitement des sarcomes d'Ewing. La protéine EWS-FLI1 est un puissant régulateur de la transcription et module l'expression de centaines voire de milliers de gènes [89]. Pour son activité de transcription elle s'associe à des protéines partenaires telles que l'ARN polymérase II, la protéine de liaison à la CREB (*C-AMP Response Element-binding protein*) et l'ARN hélicase A [292], [293]. Or il a été montré par Erkizan *et al.*, qu'il était possible d'inhiber le développement du sarcome d'Ewing en perturbant l'interaction de la protéine EWS-FLI1 avec la RNA hélicase A grâce à de petites molécules [294] (Figure 65).



**Figure 65 : Schéma de la perturbation de l'interaction de la protéine EWS-FLI1 avec l'ARN hélicase conduisant à l'inhibition de l'activité de transcription du complexe.**

Plus tard, la même équipe a réalisé un screening par la méthode de double hybride d'une série de 300 séquences peptidiques de 12 acides aminés et ont montré que 27 d'entre elles interagissaient avec EWS-FLI1. Parmi ces 27 peptides un seul s'est révélé être réellement efficace pour limiter la croissance des cellules tumorales [295]. Ce peptide, appelé *Ewing Sarcoma Antagonist Peptide 1* (ESAP1), a pour séquence TMRGKKRTRAN et présente, comme (R/W)<sub>9</sub> ou R<sub>9</sub>, un grand nombre de résidus positivement chargés.

Par conséquent, **l'hypothèse d'une interaction directe de (R/W)<sub>9</sub> avec EWS-FLI1 ou avec un de ses partenaires d'interaction susceptible de venir perturber l'activité de transcription de la protéine EWS-FLI1 mérite d'être considérée.** La présence dans la liste des protéines partenaires potentielles du polypeptide DEAD box, motif très conservé dans les RNA hélicases, va également dans le sens de cette hypothèse.

## Conclusion et perspectives

L'identification à deux reprises dans les expériences de Pull-down de la protéine EWS est un résultat très intéressant. Cependant, cette protéine n'a pas été identifiée dans les expériences de pull-down + photocross-linking indiquant que, si l'étape de photoirradiation pourrait permettre de stabiliser des interactions transitoires, elle pourrait également contribuer à masquer des interactions avec des protéines peu abondantes telles que celle avec la protéine EWS.

Il nous faut à présent valider ce résultat d'interaction avec EWS et tester notre hypothèse d'interaction avec EWS-FLI1. Pour cela nous souhaitons répéter les expériences de pull-down en



## Résultats : Pull-down et photocross-linking

utilisant des cellules 293T transfectées avec de l'ADN codant pour les protéines EWS, EWS-FLI1 et FLI1 et capables de les surexprimer. Nous travaillons actuellement avec l'Equipe du Dr. Jacques Ghysdaël de l'Institut Curie qui a le matériel et le savoir-faire nécessaire à la réalisation de cette expérience. Nous aimerions également à l'avenir, si l'interaction de (R/W)<sub>9</sub> avec la protéine EWS ou avec EWS-FLI1 est confirmée, pouvoir la caractériser *in vitro*.

Les listes de protéines partenaires d'interaction potentielles n'ont vraisemblablement pas encore pu être suffisamment exploitées. Ces résultats préliminaires pourraient être étoffés par la réalisation d'expériences de pull-down complémentaires. En particulier nous souhaiterions utiliser la version du peptide photo(R/W)<sub>9</sub> comportant des acides aminés de configuration D, non naturelle, comme contrôle négatif. En effet, il a été montré que ce peptide bien que capable de s'internaliser ne conduisait pas à un remodelage du cytosquelette d'actine dans la lignée cellulaire EF [2].





## Chapitre 7 : Evaluation de l'effet de (R/W)<sub>9</sub> sur le protéome des cellules EF par une approche de protéomique quantitative différentielle.

La capture des partenaires d'interactions du peptide photo(R/W)<sub>9</sub> ne s'étant pas avérée une tâche triviale que ce soit par photocross-linking (chapitre 5) ou purification par affinité (chapitre 6), nous avons également utilisé une approche alternative pour essayer d'obtenir des informations complémentaires sur l'origine de l'effet observé du peptide photo(R/W)<sub>9</sub> sur le phénotype des cellules EF. Une approche de protéomique quantitative différentielle, susceptible de fournir des informations sur **l'influence du peptide vecteur photo(R/W)<sub>9</sub> sur le protéome des cellules EF** a été utilisée. Puisque cette étude est réalisée sur une lignée cellulaire en culture au laboratoire, la technique de quantification SILAC (*Stable Isotope Labeling with Amino acids in cell Culture*) [296] était particulièrement adaptée. Elle permet, comme nous l'avons vu dans le chapitre 2 partie III.1 .a, de réaliser le marquage isotopique dès l'étape de culture cellulaire et ainsi de minimiser les erreurs systématiques liées à la manipulation des échantillons.

Dans cette partie, nous nous intéresserons à l'effet du peptide (R/W)<sub>9</sub> sur le protéome de cellules EF en comparant des cellules EF traitées par le peptide photo(R/W)<sub>9</sub> avec des cellules EF non traitées. Par ailleurs, nous avons pu constater **l'absence dans la littérature de données de protéomique concernant le sarcome d'Ewing**. A notre connaissance, aucune étude n'a été réalisée pour évaluer l'impact global sur le protéome de la transformation tumorale induite par l'expression de la protéine chimérique oncogénique EWS-FLI1. Par conséquent, nous avons jugé utile, dans un premier temps, de mettre en évidence les différences existant entre le protéome des cellules tumorales EF et celui des cellules non tumorales 3T3, dont elles sont issues.

### I. Mise en œuvre de la technique SILAC.

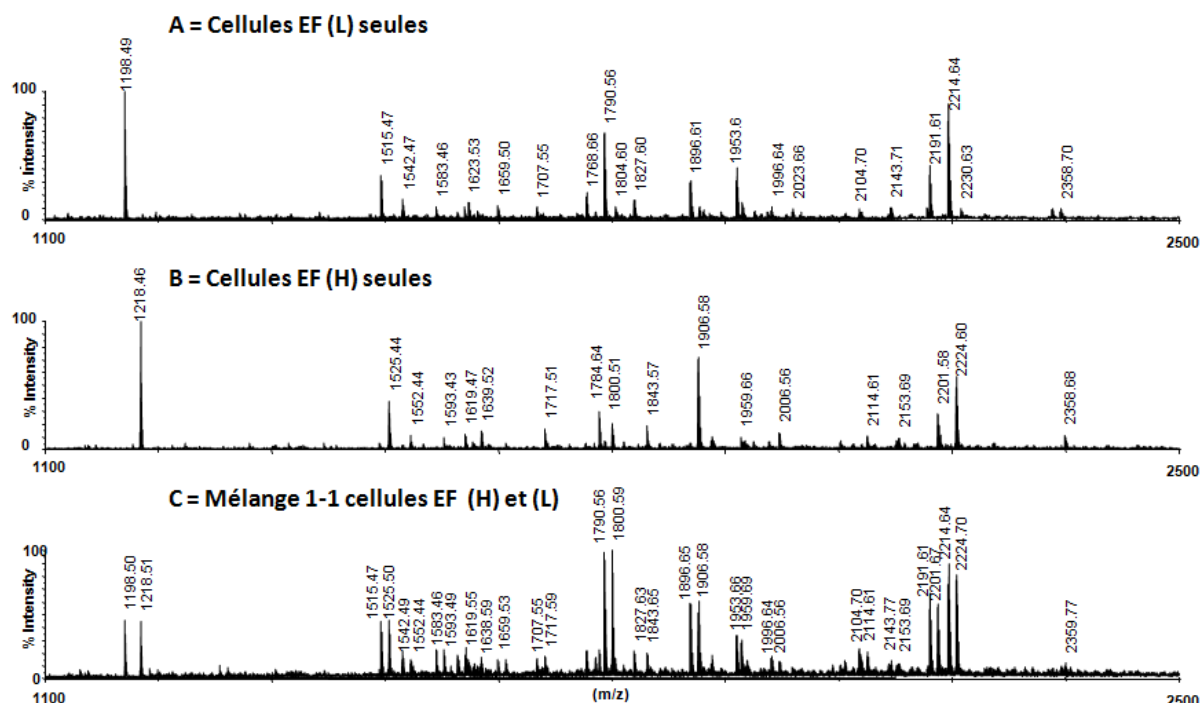
Il a été possible de mettre en place la technique SILAC très facilement et rapidement dans notre laboratoire car nous avons à la fois la possibilité de réaliser la culture cellulaire en conditions lourde et légère (présence d'acides aminés marqués isotopiquement dans les milieux de culture), le matériel pour le traitement des échantillons et l'analyse haut débit de mélanges complexes de peptides par couplage NanoLC-ESI-LTQ-OrbitrapXL (cf. partie matériel et méthodes II.2 pour détails analyses).

#### 1. Vérification du taux d'incorporation des acides aminés marqués.

La première étape consiste à vérifier que les lignées cellulaires utilisées incorporent correctement les acides aminés marqués présents dans le milieu de culture (<sup>13</sup>C<sub>6</sub> Lys et <sup>15</sup>N<sub>4</sub> Arg). Pour cela nous avons réalisé une expérience contrôle en utilisant deux populations de cellules EF identiques, l'une cultivée avec des acides aminés lourds (notés H pour Heavy) et l'autre avec des acides aminés légers (notés L pour Light) (cf. partie matériel et méthodes III.9). Cette incorporation a été vérifiée par analyse MALDI-TOF des lysats cellulaires EF obtenus pour les échantillons H et L séparément ou en mélange 1 : 1, après migration sur gel SDS-PAGE, découpage de bandes de gel de même gamme de masses moléculaires (même distance dans le gel) pour ces trois conditions et digestion à la trypsine (Figure

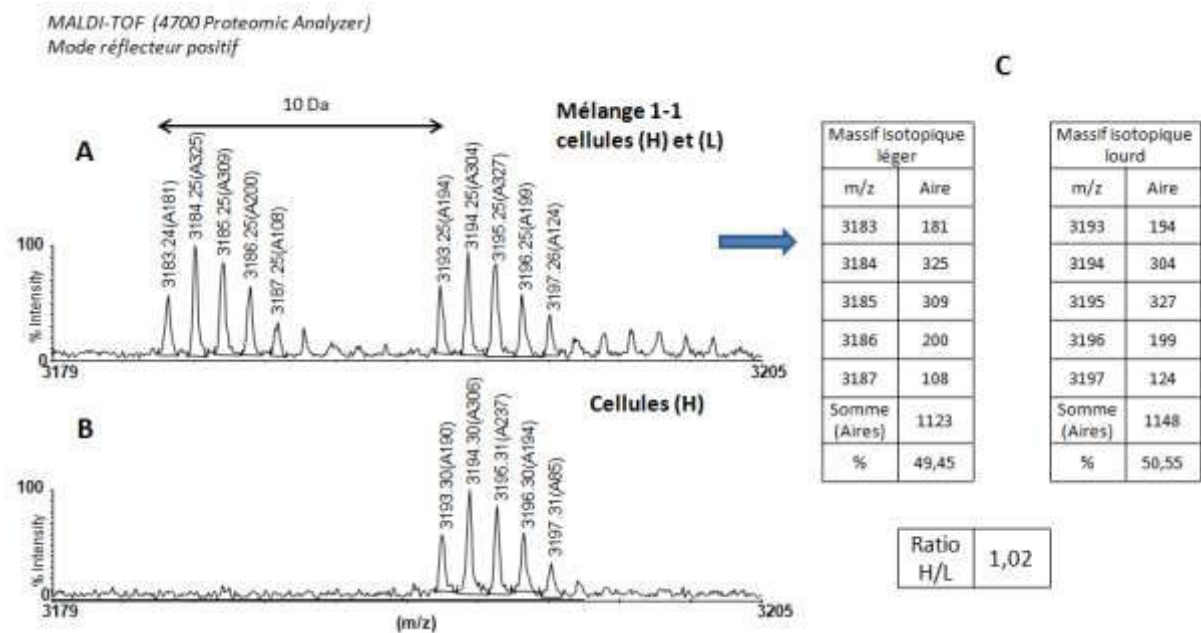
66). L'analyse MALDI-TOF a été préférée à l'analyse par nanoLC-ESI-MS/MS du fait de sa rapidité et de sa grande facilité de mise en œuvre.

Spectres MALDI-TOF (4700 Proteomic Analyzer)  
Mode réflecteur positif



**Figure 66 : Spectres MALDI-TOF pour la gamme de m/z 1100-2500 montrant le taux d'incorporation des acides aminés marqués  $^{13}\text{C}_6$  Lys et  $^{15}\text{N}_4$  Arg seuls et en mélange 1-1 des cellules (L) et (H). Le spectre A correspond à la digestion d'une bande de gel obtenue pour un dépôt de lysat cellulaire EF(L), le spectre B correspond à la digestion d'une bande de gel obtenue pour un dépôt de lysat cellulaire EF(H), Le spectre C correspond à la digestion d'une bande de gel obtenue pour un dépôt de mélange 1-1 de cellules EF (H) et (L).**

Comme observé sur le spectre **C** du mélange 1-1, chaque ion apparaît sous la forme d'un doublet d'intensité 1-1 séparé de 10 Da quand le peptide contient un résidu Arg en C-terminal et de 6 Da quand il s'agit d'un résidu Lys, en l'absence d'erreur de clivage par la trypsine. L'analyse d'un doublet particulier (Figure 67) correspondant aux versions lourdes et légères d'un même peptide permet de vérifier que (1) **l'incorporation des acides aminés marqués est totale** puisque le signal correspondant au peptide léger est complètement absent du spectre B, et que (2) **le mélange réalisé donne un ratio H/L proche de 1** comme attendu pour cette expérience contrôle.



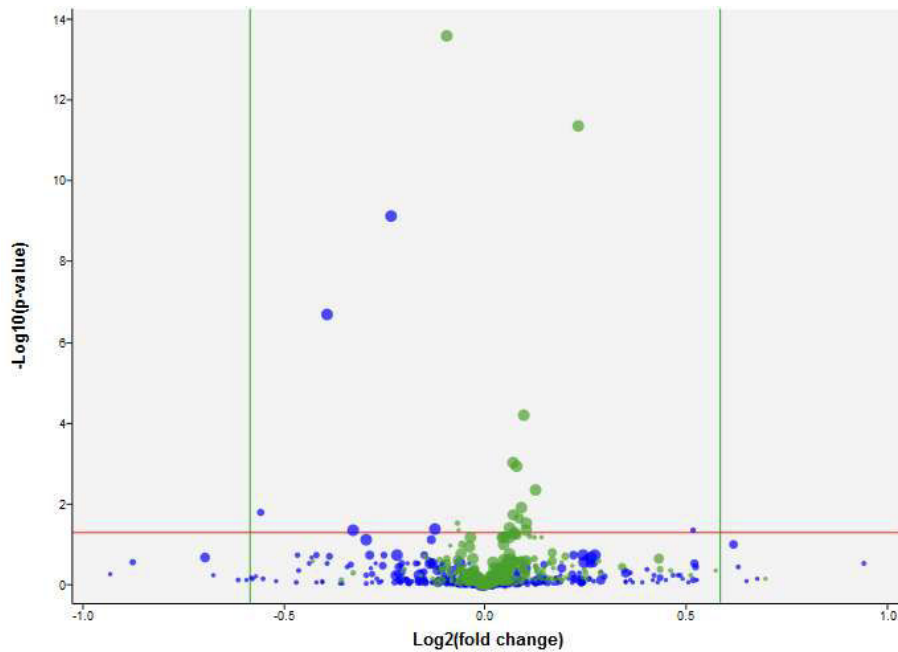
**Figure 67 : Figure illustrant la vérification du taux d'incorporation des acides aminés marqués dans les protéines.** Les parties A et B de la figure correspondent à des zooms sur les massifs isotopiques des formes mono-protonées lourdes et légères d'un peptide pour, respectivement, un mélange 1-1 de deux populations de cellules EF (H) et (L) identiques et pour des cellules EF (H) seules. La partie C de la figure indique les aires des différents pics des massifs isotopiques pour les versions (H) et (L) du peptide ainsi que la valeur du ratio H/L obtenue.

Pour réaliser le calcul du ratio (H/L) nous avons sommé l'aire sous la courbe des différents isotopes de chacun des deux massifs H et L comme indiqué Figure 67.C et exprimé cette somme en pourcentages de l'aire totale. Le ratio H/L est alors calculé en faisant le rapport de ces pourcentages. Dans ce cas, le ratio H/L a une valeur de 1,02. Nous constatons, pour cet exemple, que les deux massifs isotopiques sont séparés de 10 Da indiquant que ces deux versions du peptide contiennent une seule arginine, marquée  $^{13}\text{C}_6\text{ }^{15}\text{N}_4$  pour la version lourde.

## 2. Détermination des seuils de significativité pour la variation des ratios H/L

Une des étapes les plus délicates dans la réalisation d'expériences de protéomique quantitative différentielle est la détermination des seuils de significativité pour la variation des ratios H/L, c'est à dire les seuils qui vont permettre de décider si une protéine va être ou non incluse dans la liste des protéines significativement sur- ou sous-exprimées.

Pour la détermination de ces seuils nous avons utilisé des **échantillons contrôles correspondant au mélange de deux populations identiques de cellules EF** cultivées d'une part avec des acides aminés classiques (L), et d'autre part avec des acides aminés marqués (H). La détermination de ces valeurs seuil a été réalisée sur deux réplicats biologiques de l'expérience contrôle en analysant le protéome entier en NanoLC-ESI-MS/MS (cf. partie matériel et méthodes III.9). L'analyse statistique des données obtenues a été réalisée avec le logiciel MyProMS [297] et permet d'obtenir le graphique ci-dessous appelé volcano plot (Figure 68). Seules les protéines identifiées avec un FDR (*false discovery rate*) de 1 % et quantifiées avec un minimum de 3 peptides sont représentées.



**Figure 68 : Graphique de type volcano plot obtenu pour deux réplicats biologiques de l'expérience contrôle EF (H) / EF (L).** Chaque protéine identifiée et quantifiée est représentée par un point. A chaque réplicat correspond une couleur, ici verte ou bleue. Ce sont des échelles logarithmiques qui sont utilisées pour les deux axes. L'abscisse des points donne les valeurs  $\log_2(H/L)$  : plus un point s'éloigne de l'axe central vertical (valeur abscisse = 0) plus la variation d'expression de la protéine entre les populations H et L est importante. En abscisse cette échelle en  $\log_2$  permet de « symétriser » les ratios (H/L) de manière à ce que, par exemple, une augmentation d'un facteur 1,5 soit aussi facilement mise en évidence qu'une diminution d'un facteur 1,5.

La taille d'un point est proportionnelle au nombre de peptides ayant conduit à l'identification et à la quantification de la protéine. L'ordonnée du point renseigne sur la fiabilité du calcul du ratio H/L : plus le point est haut sur l'axe plus l'écart-type sur les valeurs de ratio H/L obtenues pour les différents peptides est faible (correspond à une valeur de p petite). L'échelle en  $\log_{10}$  permet dans ce cas de représenter une gamme de valeurs de p très étendue.

Théoriquement, les deux échantillons EF (H) et EF (L) qui sont des réplicats biologiques devraient être strictement identiques avec des volcano plot superposables (soit des ratios H/L égaux à 1 pour les différentes protéines) : les variations observées correspondent uniquement à l'erreur expérimentale. Ainsi, nous constatons que sur ce volcano plot contrôle, toutes les protéines identifiées de façon fiable avec une probabilité statistique définie par la valeur de  $p < 0,05$  (= se trouvant au-dessus de l'axe rouge horizontal) présentent une **variation de leur ratio H/L inférieure à un facteur 1,5** (= sont comprises entre les deux barres vertes verticales).

Ce sont par conséquent ces valeurs ( $p < 0,05$  et  $H/L > 1,5$  (surexprimées) ou  $p < 0,05$  et  $H/L < 1/1,5 = 0,67$  (sous-exprimées)) qui vont être choisies comme valeurs seuil pour définir les zones correspondant à des protéines significativement sous- ou surexprimées dans les expériences de protéomique différentielle qui vont suivre.

**Seuils de significativité utilisés dans la suite de l'étude :**

Les protéines (= points) présentes dans le rectangle en haut à gauche du volcano plot ont une valeur  **$p < 0,05$  et un ratio  $H/L < 1/1,5 = 0,67$**  et correspondent à des protéines **significativement sous-exprimées** dans la population (H) par rapport à la population (L).

Les protéines (= points) présentes dans le rectangle en haut à droite du volcano plot ont une valeur  **$p < 0,05$  et un ratio  $H/L > 1,5$**  et correspondent à des protéines surexprimées dans la population (H) par rapport à la population (L).

**Remarque :** Le choix d'une valeur de  $p < 0,05$  nous amène à exclure un grand nombre de protéines identifiées mais quantifiées de manière peu robuste. Pour les expériences contrôles réalisées précédemment chaque bande a été analysée une seule fois et un nombre relativement restreint de peptides a ainsi été identifié pour chacune des protéines ce qui nous conduit à avoir la grande majorité des protéines avec une p value élevée. Pour les expériences de protéomique différentielle qui suivent, chaque bande a été analysée au minimum deux fois permettant d'améliorer la statistique sur la quantification des protéines et ainsi d'en considérer davantage en termes de variation des ratios.

Une fois les seuils de significativité établis, la comparaison des protéomes des cellules tumorales EF avec leur équivalent non-tumoral 3T3 a pu être effectuée.

## II. Comparaison des protéomes des cellules EF et 3T3.

Pour cette expérience, nous avons comparé une population de cellules EF cultivées en présence d'acides aminés lourds (EF (H)) avec une population de cellules 3T3 cultivées en présence d'acides aminés légers (3T3 (L)), toutes choses égales par ailleurs.

Trois réplicats biologiques ont été réalisés. Conformément aux seuils de significativité définis précédemment nous n'avons conservé que les protéines présentant une variation du ratio H/L supérieur à un facteur 1,5 ainsi qu'une valeur  $p < 0,05$  sur la détermination du ratio H/L. De plus, nous avons choisi de ne conserver que les **protéines visibles au moins dans deux des trois réplicats biologiques**. Le volcano plot obtenu pour cette expérience est présenté Figure 69.

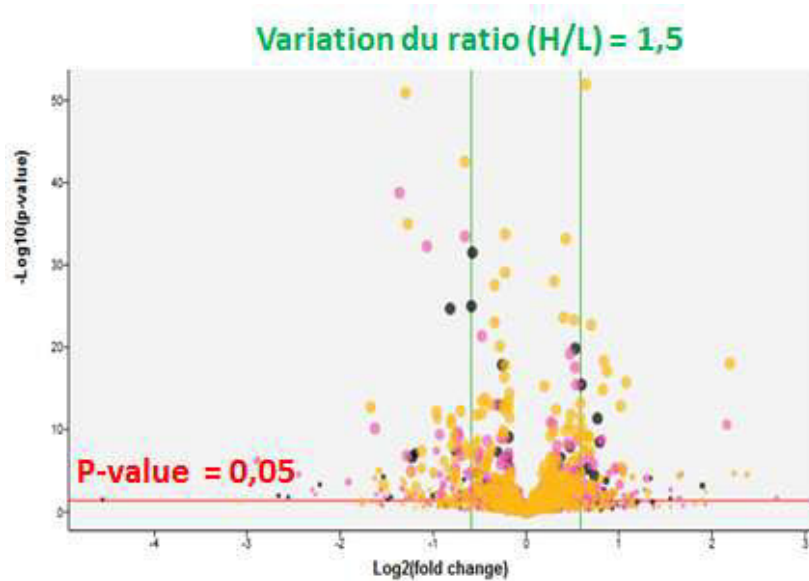


Figure 69 : Graphique de type volcano plot pour la comparaison du protéome des cellules EF (H) et des cellules 3T3 (L). A chacun des trois réplicats biologiques réalisés correspond une couleur, ici orange, rose et noire.

L'application de ces filtres sur les 1700 protéines identifiées et quantifiées pour les trois réplicats nous a permis d'obtenir une liste de **106 protéines présentant une variation significative de leur niveau d'expression** entre les cellules tumorales EF et leur équivalent non-tumoral 3T3. Parmi elles, 50 sont surexprimées et 56 sous-exprimées dans les cellules EF par rapport aux cellules 3T3. (Listes de protéines significativement sur- et sous exprimées présentées en annexes N°6)

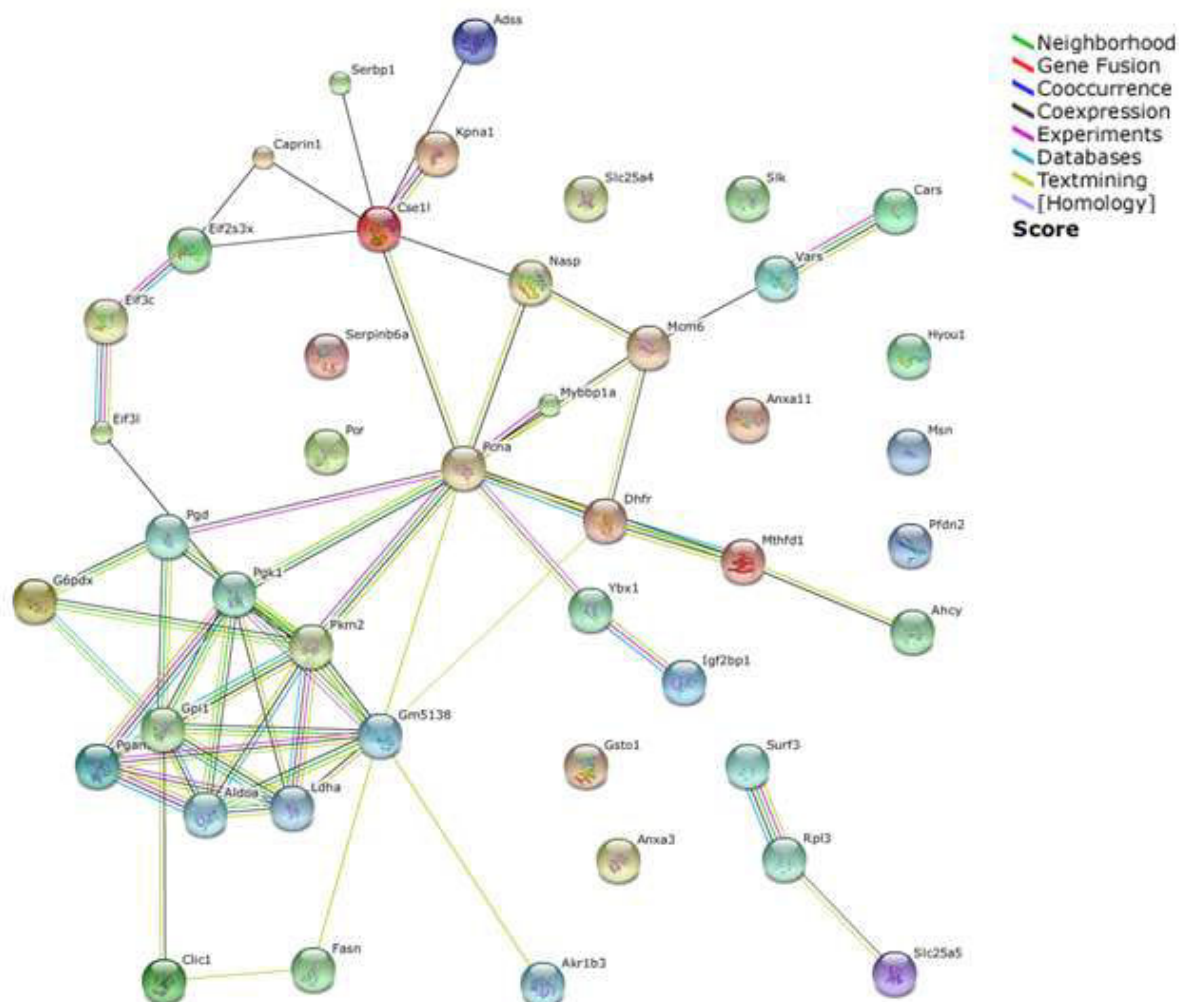
Afin de donner un sens biologique à ces protéines et mettre en évidence des protéines susceptibles d'expliquer le changement de phénotype accompagnant la transformation tumorale pour la protéine EWS-FLI1, nous avons fait appel à la base de données d'interactions protéine-protéine STRING (*Search Tool for the Retrieval of Interacting Genes/Proteins*)[195],[298]. STRING permet également de faire de l'enrichissement de termes de *gene ontology* (cf. chapitre 2, partie III.2).

Avec cet outil, nous avons mis en évidence des fonctions biologiques significativement enrichies parmi les protéines sur- ou sous-exprimées obtenues.



## 1. Protéines significativement surexprimées dans les cellules EF par rapport aux cellules 3T3.

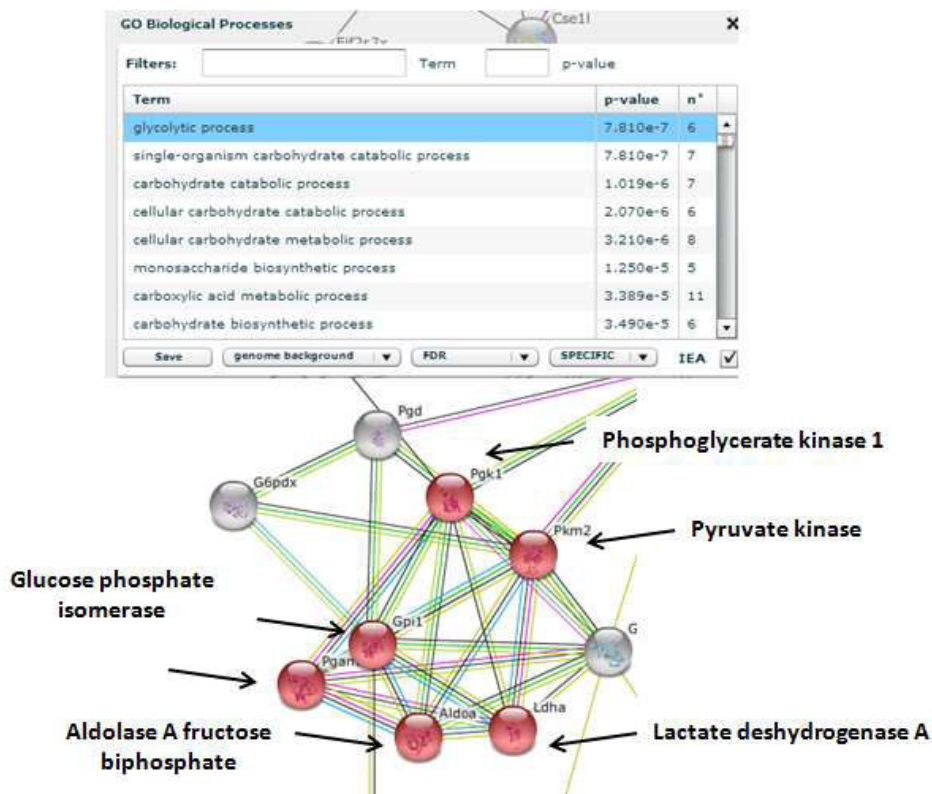
Les protéines significativement surexprimées dans les cellules EF par rapport aux cellules 3T3 sont représentées ci-dessous sous forme d'un réseau d'interaction protéine-protéine (Figure 70).



**Figure 70 : Réseau d'interactions protéine-protéine généré sur le logiciel STRING à partir de la liste de protéines surexprimées dans les cellules EF par rapport aux cellules 3T3.** Chaque sphère correspond à une protéine et chaque trait correspond à un lien entre deux protéines dans la base de données. La couleur des traits fait référence au type de lien existant entre ces protéines comme indiqué dans la légende (en haut à droite).

### a. Augmentation de la quantité des protéines impliquées la glycolyse.

Le nœud d'interaction le plus dense de ce réseau correspond à des protéines directement liées au processus de glycolyse. L'enrichissement en termes de *gene ontology* des protéines de ce réseau en fonction des processus biologiques lui donne la plus petite valeur p ( $p \approx 10^{-6}$ ) (Figure 71).



**Figure 71 : Nœud du réseau d'interaction correspondant au processus de glycolyse.** Les sphères colorées en rouge correspondent aux protéines qui sont directement impliquées dans le processus de glycolyse et qui ont conduit à l'enrichissement de ce terme de *gene ontology*.

Cette augmentation de la glycolyse correspond à une propriété quasi-universelle des cellules cancéreuses primaires ou métastatiques. En effet, en plus de l'ATP nécessaire pour maintenir les processus cellulaires normaux, les cellules cancéreuses ont besoin d'énergie supplémentaire pour satisfaire leurs besoins élevés, liés à leur division rapide.

Dans des conditions aérobies classiques, une cellule normale va métaboliser le glucose afin de produire de l'ATP, source d'énergie de la cellule. Ce processus fait intervenir 1) la glycolyse qui transforme une molécule de glucose en deux molécules de pyruvate puis 2) le cycle de Krebs qui va permettre de dégrader complètement le pyruvate en  $\text{CO}_2$  et  $\text{H}_2\text{O}$  avec production d'ATP par des étapes successives de phosphorylation oxydative réalisées dans les mitochondries. Ce processus complet est relativement long mais très efficace pour la production d'ATP (36 à 38 molécules d'ATP produites par molécule de glucose).

Dans les mêmes conditions, une cellule cancéreuse va être capable de contourner ce long processus pour parer au plus urgent en n'utilisant que la première étape du processus, à savoir la glycolyse. Ce stratagème permet à la cellule cancéreuse de produire de l'ATP plus rapidement mais de façon beaucoup moins rentable puisque dans ce cas, seules 2 molécules d'ATP sont produites directement lors de la glycolyse et 6 supplémentaires par phosphorylation oxydative du NADH produit lors de la glycolyse. Ce phénotype métabolique est connu sous le nom **d'effet Warburg** [299].

Ainsi, pour une production d'ATP à la fois plus rapide et plus importante, la cellule cancéreuse doit consommer rapidement une grande quantité de glucose, ce qui explique la surexpression des

protéines impliquées dans le processus de glycolyse pour faire face à cette demande accrue et rapide d'énergie [300].

### b. Augmentation des protéines impliquées dans la biosynthèse

Parmi les 50 protéines significativement surexprimées, 17 sont impliquées dans les processus de biosynthèse (biosynthèse de monosaccharides, de nucléotides, d'acides aminés...).

Les cellules tumorales qui ont une vitesse de croissance augmentée ont, en effet, besoin de produire davantage de "briques élémentaires" pour **garantir leur capacité de prolifération** [301]. Cette observation est en accord avec l'augmentation des protéines de la glycolyse puisque davantage d'énergie va être consommée.

### c. Protéines oncogéniques

Nous pouvons noter également la surexpression de différentes protéines connues pour leur rôle important dans le processus d'oncogénèse. Parmi elles, la moesin dont la surexpression est nécessaire à la transition épithélio-mésenchymateuse des cellules cancéreuses conduisant à la formation de métastases [302].

Une autre protéine importante est l'*insulin like growth factor 2* (IGF2) qui est en fait un peptide mitogène de 7,5 kDa surexprimé dans différents cancers et généralement associé à un mauvais pronostic [303].

### d. Protéine impliquée dans le remodelage de l'actine

Parmi les protéines surexprimées dans les cellules EF par rapport aux 3T3, il semblerait qu'il y ait au moins une protéine correspondant au phénotype particulier observé pour les cellules EF. En effet, il a été montré que la STE20 like serine/threonine protein kinase (SLK) jouait un rôle dans le remodelage du cytosquelette [304]. De plus, il a été observé qu'une surexpression de SLK activée induisait une rétraction des cellules vis-à-vis de leur substrat et inhibait le déplacement des cellules. Il semblerait que cet effet soit lié à la dissolution des fibres des stress par SLK [305].

## 2. Protéines significativement sous-exprimées dans les cellules EF par rapport aux cellules 3T3.

Pour les protéines sous-exprimées dans les cellules EF par rapport aux cellules 3T3, nous avons procédé de la même manière en nous appuyant sur le réseau d'interaction construit à partir de la liste de protéines *via* le logiciel STRING (réseau d'interaction présenté en annexes N° 7).

### a. Diminution des processus d'oxydo-réduction et de respiration

Parmi les protéines sous-exprimées, 13 sont directement impliquées dans des processus d'oxydo-réduction.

Les cellules cancéreuses, comparées aux cellules normales, se trouvent dans une **situation de stress oxydant** liée à une activité métabolique accrue réalisée en limitant le fonctionnement des mitochondries et les processus de respiration [306]. Ce stress oxydant se traduit par l'augmentation du niveau des espèces réactives à base d'oxygène (ROS *reactive oxygen species*). Les ROS incluent des

radicaux libres comme les superoxydes ( $O_2^{\bullet-}$ ) et les radicaux hydroxyles ( $HO^{\bullet}$ ) ainsi que des molécules non radicalaires telles que  $H_2O_2$ . Il a été montré que cette **augmentation du niveau de ROS** peut avoir de nombreuses conséquences bénéfiques du point de vue de la cellule cancéreuse. En effet, il apparaît que les ROS stimulent la prolifération cellulaire et entraînent des **mutations et des instabilités génétiques** pouvant conduire, par exemple, au **développement de résistances** aux traitements anticancéreux [307]. Le stress oxydant pourrait ainsi être favorable au développement des tumeurs. Cependant, les mécanismes responsables de ce stress oxydant observé dans la plupart des cancers ne sont pas précisément connus. En particulier, il est difficile de savoir quel est le niveau critique de ROS à partir duquel on passe d'une situation propice au développement du cancer à celle d'une menace pour la survie de ces cellules.

Quoi qu'il en soit, la modification des niveaux d'expression des protéines impliquées dans les processus d'oxydo-réduction constitue une caractéristique commune des cellules cancéreuses. Parmi ces protéines, beaucoup de protéines aux propriétés antioxydantes sont identifiées dans nos expériences avec un niveau d'expression diminué dans les cellules tumorales EF par rapport aux cellules 3T3 : C'est le cas des superoxide dismutases (SOD) [Cu-Zn] et [Mn], des Glutathione-S-transférases Mu1 et Mu2, des NADPH isocitrate déshydrogénase cytoplasmique et mitochondriale ou encore de la peroxiredoxine-2. Dans beaucoup de cancers, les niveaux d'expression des SOD sont au contraire augmentés, vraisemblablement pour éviter d'atteindre un niveau cytotoxique de ROS [308], [309]. Cependant dans d'autres cancers le niveau de SOD est diminué [310], [311] comme c'est le cas dans notre étude (diminution d'un facteur 1,5-1,6). Cela suggère que les différents types de cancers régulent de manière différente leur niveau de ROS et que, pour le sarcome d'Ewing, l'obtention d'un niveau satisfaisant de ROS passerait par la diminution du niveau d'expression des protéines anti-oxydantes.

Par contre, il est important de noter la surexpression de la *hypoxia up-regulated protein 1* dans les cellules EF par rapport aux cellules 3T3 qui indique que les cellules EF sont également dans des conditions de **stress hypoxiques**, c'est à dire de déficit en oxygène. Différentes études [312], [313] suggèrent que les cellules cancéreuses se placent volontairement dans ces conditions de manière à s'adapter à l'absence d'oxygène. En effet l'oxygène fait défaut aux tumeurs en développement dans un tissu jusqu'à la mise en place de nouveaux réseaux sanguins nécessaires à leur oxygénation (angiogénèse).

### e. Sous-expression de protéines de liaison de l'actine.

Le cytosquelette d'actine joue un rôle crucial dans la vie de la cellule (division, migration, transmission des signaux...) et les mécanismes de régulation de la dynamique de l'actine reposent sur l'intervention de très nombreuses protéines de liaison de l'actine (*cf.* chapitre 1. Partie II .3). Il est facile d'imaginer que ces protéines peuvent subir des modifications d'expression lors de la transformation tumorale. Cependant ces modifications des protéines du cytosquelette correspondant au changement de phénotype induit par la protéine oncogénique EWS-FLI1 n'ont pas encore, à notre connaissance, été décrites dans la littérature.

L'examen attentif des listes de protéines significativement sous-exprimées obtenues pour la comparaison des cellules EF et 3T3, a permis d'identifier un total de 12 protéines connues pour leur

liaison et leur implication dans la dynamique de l'actine. Ces protéines sont présentées dans le Tableau 13.

**Tableau 13 : Récapitulatif des protéines de liaisons de l'actine figurant dans la liste des protéines significativement sous-exprimées dans les cellules tumorales EF par rapport aux cellules non-tumorales 3T3.**

Protéines de liaison de l'actine	Fonction(s)	Références
LIM and SH3 domain protein 1 (LASP-1)	Régulation des fonctions cellulaires associées avec les réarrangements <b>du cytosquelette à la membrane</b> . Présente dans les points d'adhésion focaux et associée notamment à la zyxine.	[314]
Filamine	<b>Cross-linker les filaments d'actine</b> pour la formation de réseaux orthogonaux d'actine	[46] + cf. chapitre 1. partie II.3.b.
Destrine (ou Actin Depolymerizing Factor ADF)	<b>Séquestre les filaments d'actine</b> et se lie aux monomères d'actine	[315]
Protéine AHNAK 1 et 2	<b>Formation de faisceaux d'actine</b>	[316]
Gelsoline	Protéine de <b>ségrégation</b> et de coiffe	[38] + cf. chapitre 1 partie II.3.a
Myosine (myosin regulatory light chain 12B (MYL12B) + myosin-9+ myosin light polypeptide 6)	Constituants du complexe Myosine II qui lie l'actine et joue un rôle central dans l'adhésion cellulaire, la migration et la division.  Une sous-régulation <i>des myosin regulatory light chains A ou B (MYL12B/12A)</i> induit un changement remarquable de morphologie cellulaire et la disparition des fibres de stress	[317]
Tropomyosine	Protéine de <b>stabilisation</b> latérale des filaments	[41] + cf. chapitre 1 partie II.3.a
Caldesmone	Protéine de <b>stabilisation</b> latérale des filaments	[43] + cf. chapitre 1 partie II.3.a
Calponine-3	Joue un rôle dans la <b>formation des fibres de stress</b>	[318]

A l'exception de la destrine et de la gelsoline qui sont plutôt connues pour maintenir l'actine sous forme monomérique, les protéines de liaison de l'actine sous-exprimées dans les cellules EF, ont toutes un rôle soit dans la stabilisation des filaments d'actine (tropomyosine, caldesmone), soit dans leur assemblage en réseaux ou en faisceaux (calponine-3, protéine AHNAK 1 et 2, filamine, sous unités de la myosine II), ou encore dans leur ancrage à la membrane (LASP-1).

La comparaison des cellules EF et 3T3 réalisée indique par conséquent que la transformation tumorale conduit à une **réduction significative de l'expression d'un nombre important de protéines clés de liaison de l'actine**. Ce résultat va à l'encontre de ce que l'on peut attendre pour des cellules cancéreuses qui sont amenées à migrer par formation de filipodes, de lamellipodes, etc... (cf. chapitre 1 partie III.1).

Par contre, cette observation va dans le sens de l'étude de Chaturvedi *et al.* [93] qui a mis en évidence le comportement tout à fait particulier des sarcomes d'Ewing par rapport à la majorité des cellules cancéreuses et qui faisait l'hypothèse d'un mode de dissémination passif de ces cellules.

Ainsi il semblerait que le phénotype des cellules EF caractérisé par un cytosquelette désorganisé dépourvu de fibres de stress, une forme ronde et l'absence de points d'adhésion focaux, soit le résultat d'une sous-expression des protéines de liaison de l'actine.

### III. Effet du peptide vecteur photo(R/W)<sub>9</sub> sur le protéome des cellules EF.

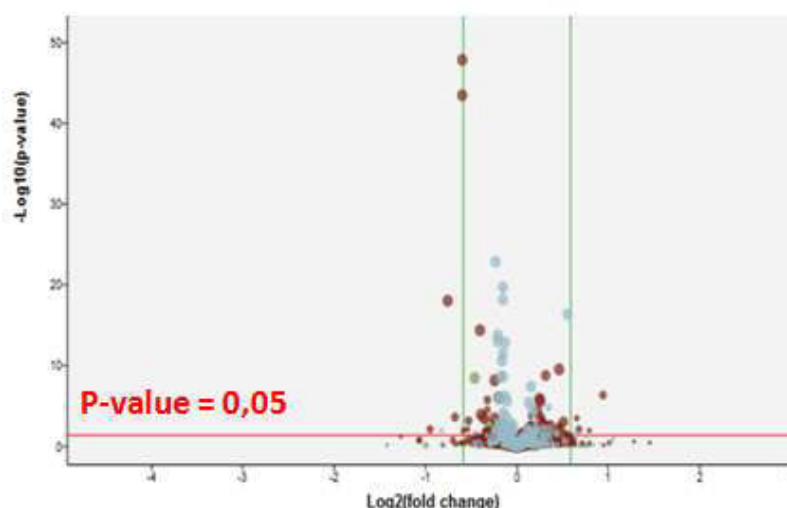
Après avoir mis en évidence des changements d'expression des protéines de liaison de l'actine pouvant expliquer le changement de phénotype entraîné par la protéine oncogénique EWS-FLI1, nous nous sommes intéressés aux modifications d'expression de protéines liées à la présence du peptide vecteur photo(R/W)<sub>9</sub>. Nous souhaitons essayer de corréliser les variations du protéome observées entre les cellules EF et 3T3 avec celles liées à l'internalisation du peptide photo(R/W)<sub>9</sub>.

Pour cela nous avons comparé le protéome des cellules EF(H) non traitées avec celui des cellules EF(L) traitées une nuit avec une concentration extracellulaire de 5  $\mu$ M en peptide photo(R/W)<sub>9</sub>. Dans le chapitre 5, partie I.2. , nous avons vu qu'au bout de 75 min la quantité de peptide photo(R/W)<sub>9</sub> internalisé dans 10<sup>6</sup> cellules était de l'ordre de 3 pmol ce qui correspond à une dizaine de millions de molécules de peptide photo(R/W)<sub>9</sub> internalisées par cellules (le volume de 10<sup>6</sup> cellules étant d'environ 1,5  $\mu$ L cela correspond à une concentration intracellulaire de l'ordre de 2  $\mu$ M).

De la même manière que pour la comparaison des cellules EF et 3T3, trois réplicats biologiques de l'expérience ont été réalisés. A nouveau, les seuils de significativité définis grâce aux expériences contrôles (une variation du ratio H/L supérieure à un facteur 1,5 ainsi qu'une valeur de  $p < 0,05$  sur la détermination du ratio H/L) ainsi que le critère de présence dans au moins deux des trois réplicats ont été utilisés pour filtrer les résultats. Le volcano plot obtenu dans ce cas est présenté Figure 72.



Variation du ratio (H/L) = 1,5



**Figure 72 : Graphique de type volcano plot pour la comparaison du protéome des cellules EF (H) et des cellules EF (L) incubées avec le peptide photo(R/W)<sub>9</sub> (18 h, 5 μM).** A chacun des trois réplicats biologiques réalisés correspond une couleur, ici grise, brune et verte.

Dans ce cas, aucune protéine, parmi les 1745 protéines identifiées (1 % FDR) et quantifiées (avec 3 peptides minimum), n'est significativement modifiée de manière reproductible.

Par conséquent, il apparaît qu'au bout de 18 h de traitement avec le peptide photo(R/W)<sub>9</sub>, le changement de phénotype observé en immunofluorescence ne se traduit pas par une modification importante et quantifiable d'expression des protéines.

Cette expérience donne ainsi lieu à la question suivante : « **Comment se fait-il qu'un changement de phénotype des cellules EF engendré par le peptide photo(R/W)<sub>9</sub> ne soit pas visible au niveau du protéome de ces cellules ?** ». A notre connaissance ce genre de question n'a jamais été réellement posée dans la littérature.

Pour essayer de répondre à cette question les hypothèses suivantes ont été formulées :

- Il est possible que nous ne soyons pas en mesure de mettre en évidence ces changements. En effet, nous avons noté lors des expériences d'immunofluorescence, l'hétérogénéité des phénotypes existant dans une population de cellules traitées (cf. présentation du projet de recherche : environ 45 % de cellules N (non-tumorales), 35 % de cellules ND (non-déterminées) et 20 % de cellules T (tumorales)). Par conséquent, nous pouvons imaginer que les changements d'expression correspondant à ce changement de phénotype soient moyennés avec les niveaux d'expression de ces protéines dans des cellules qui ne seraient pas modifiées par le traitement avec le peptide photo(R/W)<sub>9</sub>.
- Nous pouvons également imaginer que l'obtention d'un phénotype se rapprochant de celui des cellules 3T3 non-tumorales à l'issue du traitement avec photo(R/W)<sub>9</sub>, soit le résultat de cascades d'évènements moléculaires différentes de celles induites par la transformation par EWS-FLI1. Le peptide photo(R/W)<sub>9</sub> pourrait, par exemple, **plutôt induire des modifications post-traductionnelles que des variations d'expression des protéines**. Il est également envisageable que les niveaux d'expression d'autres protéines que celles que nous avons pu

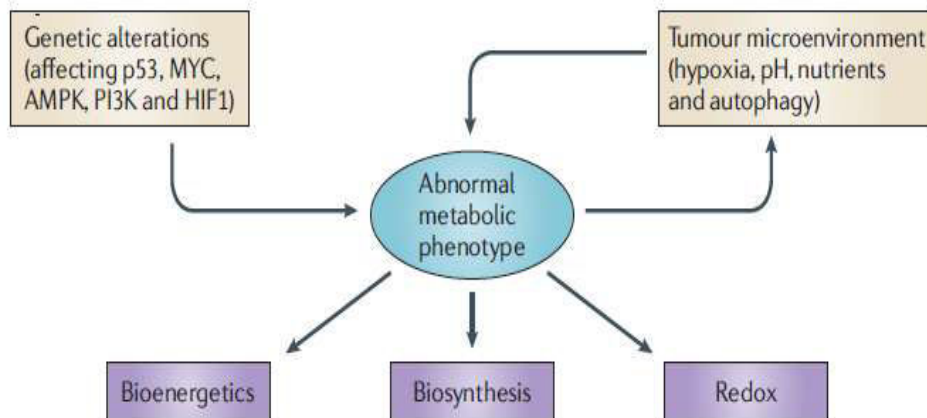


identifier dans notre expérience soient modifiés (seulement environ 1700 protéines identifiées sur un total de l'ordre de 10 000 à 12 000 protéines recensées [319]). Pour pouvoir répondre à cette question, il nous faudrait être en mesure d'identifier un plus grand nombre de peptides de façon (1) à identifier de nouvelles protéines et (2) à obtenir une meilleure statistique pour la quantification des protéines déjà identifiées et ainsi pouvoir considérer davantage de valeurs de ratio (H/L). Pour cela nous pourrions augmenter le nombre de bandes de gels utilisées pour le préfractionnement, envisager un préfractionnement par focalisation isoélectrique [320] ou en chromatographie échangeuse de cations (SCX) [321]. Un préfractionnement au niveau cellulaire par ultracentrifugation différentielle est également envisageable (cf. partie matériel et méthodes III.7).

- Il est possible qu'une durée d'incubation de 18 h ne soit pas adaptée pour mettre en évidence les modifications du protéome. Il serait ainsi intéressant de réaliser l'expérience pour des temps plus courts (ex : 2 h) ou plus longs (ex : quelques jours).

### Conclusion :

La comparaison des protéomes de cellules tumorales EF et des cellules non-tumorales 3T3 dont elles sont issues, nous a permis de mettre en évidence des variations d'expression de protéines caractéristiques de différents processus connus pour être modifiés lors de la transformation tumorale. Les **modifications du métabolisme** des cellules tumorales sont récapitulées Figure 73 [322].



**Figure 73 : Modifications du métabolisme caractéristiques de la transformation tumorale** (figure tirée de [322])

Au-delà du métabolisme, **des changements sont attendus au niveau du cytosquelette** lors de la transformation tumorale. Classiquement les cellules tumorales qui cherchent à se disséminer vont avoir tendance à surexprimer de nombreuses protéines de liaison de l'actine et à renforcer leur cytosquelette pour migrer. Les cellules tumorales EF, modèle cellulaire du sarcome d'Ewing, semblent différer sur ce point des cellules tumorales classiques puisque **l'expression de beaucoup de protéines de liaison de l'actine est diminuée** à l'exception de la STE20 like serine/threonine protein kinase qui est connue pour empêcher la migration des cellules [305]. Ainsi, il apparaît que

conformément à l'hypothèse de Chaturvedi et al. [93] , tout soit mis en œuvre pour limiter l'ancrage de ces cellules et leur migration active, de façon à **privilégier un mode de dissémination passif**.

Alors que ces expériences de protéomique quantitative se sont avérées utiles pour comprendre comment la transformation tumorale induite par EWS-FLI1 affecte le protéome des cellules fibroblastiques, nous n'avons, pour le moment, pas pu mettre en évidence de changements significatifs du protéome des cellules EF induits par le peptide photo(R/W)<sub>9</sub> malgré les modifications observées au niveau phénotypique.



## Conclusion

La compréhension des mécanismes d'action des CPPs présentant une activité biologique intrinsèque est un prérequis à leur utilisation à des fins thérapeutiques comme agents de pharmacomodulation et/ou de vectorisation. Cette étude s'est concentrée sur la **compréhension de l'effet moléculaire du CPP (R/W)<sub>9</sub> sur des cellules modèle du sarcome d'Ewing (cellules EF)**. Toutes les approches développées dans ce travail pourront être étendues à l'étude d'autres CPPs présentant des propriétés biologiques distinctes, de même qu'à d'autres systèmes d'interactions protéine-protéine.

Ce projet de recherche nous a amené, tout d'abord, à évaluer la méthode de **cross-linking chimique couplée à la spectrométrie de masse** pour l'étude de complexes CPP-protéine. A l'issue d'optimisations des conditions de réaction et d'analyse et après avoir mis au point des outils adaptés pour l'interprétation des données, nous avons pu obtenir à la fois des informations sur la composition des complexes entiers CPP(s)-protéine cross-linkés et sur les zones d'interaction. Cette approche **in vitro** nous a permis d'identifier des zones d'interactions de l'actine avec le CPP (R/W)<sub>9</sub> connues pour jouer un rôle crucial dans sa dynamique.

Cependant étant donné la multiplicité des zones d'interaction du CPP (R/W)<sub>9</sub> avec l'actine, la question de la spécificité de cette interaction s'est posée et nous a incité à entreprendre des expériences dans un contexte plus proche de la réalité c'est-à-dire sur cellules entières (**in cellulo**) ou sur **lysats cellulaires obtenus en conditions non-dénaturantes**. Pour cela, il a été nécessaire de s'affranchir des cross-linkers généralement non perméables aux membranes biologiques et s'hydrolysant rapidement en milieu aqueux. Ainsi, pour la capture des partenaires d'interaction du peptide (R/W)<sub>9</sub>, nous avons développé une méthode de **photocross-linking couplé à la spectrométrie de masse**. Dans ce cas, il est possible de laisser le peptide vecteur entrer dans les cellules, interagir avec ses partenaires et de seulement ensuite, au moment voulu par l'expérimentateur, réaliser la capture des partenaires du CPP par photoirradiation des cellules entières.

Nous avons pour cela conçu et fait synthétiser une version photoactivable du peptide (R/W)<sub>9</sub>. Cette version du peptide nommée photo(R/W)<sub>9</sub>, comprenant une benzophénone pour la réaction de photocross-linking et une étiquette biotine pour la purification des complexes cross-linkés, a été validée d'un point de vue biologique afin de s'assurer que l'introduction de ces nouvelles fonctions n'altérerait ni sa capacité d'internalisation, ni sa capacité de remodelage de l'actine. La cytotoxicité du peptide photo(R/W)<sub>9</sub> a également été évaluée afin de s'assurer qu'elle était compatible avec la réalisation des expériences sur cellules entières. A l'issue de cette **validation biologique**, la réaction de photocross-linking a été réalisée *in vitro* avec le système modèle BSA-photo(R/W)<sub>9</sub>. L'**absence de logiciels dédiés à l'interprétation des spectres MS et MS/MS des espèces photocross-linkées** nous a incité à prolonger la collaboration avec l'équipe du Dr X. Du pour développer une nouvelle version du logiciel Xlink-Identifier, jusqu'alors limité à l'interprétation des analyses de milieux de cross-linking chimique [323]. Cette nouvelle version, **Xlink-Identifier 2.0**, en cours de finalisation, doit permettre d'identifier les partenaires d'interaction d'un peptide porteur d'une sonde photoactivable et de déterminer les résidus impliqués dans les réactions de photocross-linking à partir des **spectres CID et ETD des espèces photocross-linkées**. Ce travail nous a conduits à nous intéresser de près aux

spectres de fragmentation ETD et CID pour en tirer le maximum d'informations utiles à l'identification des sites de photomarquage. En effet, seule une partie de l'information contenue dans ces spectres est actuellement prise en compte dans les recherches automatisées. Il reste encore des optimisations à réaliser afin de tenir compte de la totalité des informations pertinentes.

De manière plus générale, les logiciels dédiés à des applications biologiques particulières (comme ici le photomarquage) font encore souvent défaut bien qu'incontournables pour traiter les quantités énormes d'informations générées par les études de protéomique. L'évolution de ces moteurs de recherche passe irrémédiablement par une étroite collaboration entre les bioinformaticiens et les bio-analyticiens et une harmonisation de leurs façons de penser et de conceptualiser les choses est nécessaire. Cette collaboration avec l'équipe du Dr X. Du a été très enrichissante de ce point de vue.

Nous avons par ailleurs travaillé à la mise au point d'un protocole pour la réalisation d'expériences de **photocross-linking in cellulo**. Cela nous a permis de vérifier qu'il y avait bien formation de complexes photocross-linkés lors de l'étape de photoirradiation sur cellules entières, mais cela nous a également permis de prendre conscience d'un certain nombre de difficultés. Tout d'abord, nous avons pu constater l'impossibilité de la capture par affinité des complexes photocross-linkés entiers : la sonde biotine est vraisemblablement rendue peu, voire inaccessible après la liaison covalente du peptide photo(R/W)<sub>9</sub> avec des protéines partenaires. Par ailleurs, ces expériences ont mis en évidence la grande multiplicité des partenaires d'interaction de ce peptide qui est un concentré de « points chauds » [290] avec à la fois des résidus arginine chargés positivement et des résidus tryptophane hydrophobes. Ce grand nombre de partenaires contribue à diluer fortement le signal MS pour les espèces photocross-linkées analysées après digestion enzymatique (trypsine) et purification d'affinité. Ainsi, à l'heure actuelle, la méthode développée n'a pas permis d'identifier des partenaires d'interaction du peptide photo(R/W)<sub>9</sub>. Ces méthodologies appliquées dans le cas particulier de notre étude conservent un caractère tout à fait général et pourront être utilisées pour l'étude de partenaires d'interaction d'autres peptides ou petites molécules présentant des caractéristiques physicochimiques moins propices aux interactions non spécifiques.

L'impossible capture sur billes magnétiques de streptavidine des complexes photocross-linkés entiers nous a amenés à nous tourner vers la méthode de **pull-down**. Cette technique couplée à la spectrométrie de masse et associée au logiciel de traitement de données SAINT, a permis d'identifier **deux partenaires d'interaction du peptide photo(R/W)<sub>9</sub> pertinents d'un point de vue biologique: la protéine EWS et le polypeptide DEAD box**, motif commun aux RNA hélicases. L'identification de ces partenaires potentiels suggère la possibilité d'une **perturbation de l'activité de transcription de la protéine EWS-FLI1 par le CPP (R/W)<sub>9</sub>**. Cependant cette hypothèse, construite sur des résultats préliminaires, demande à être confirmée par différentes expériences. Certaines reposant sur la surexpression des protéines EWS, EWS-FLI1 et FLI1 sont actuellement en cours.

Finalement, nous avons choisi d'évaluer l'influence du CPP (R/W)<sub>9</sub> sur le protéome des cellules EF. Avant cela et étant donné le déficit des données de protéomique à ce sujet, nous nous sommes intéressés à l'effet de la transformation tumorale induite par la protéine oncogénique EWS-FLI1 sur le protéome. Pour cela nous avons utilisé une approche de **protéomique quantitative** reposant sur le marquage métabolique **SILAC** et avons comparé les **protéomes des cellules tumorales EF et des cellules non-tumorales 3T3**. Cette expérience a révélé la sous-régulation lors de la transformation

tumorale de l'expression de nombreuses protéines de liaison de l'actine jouant un rôle crucial dans la dynamique du cytosquelette. Ces résultats sont en accord avec le phénotype des cellules EF et appuient l'hypothèse formulée par Chaturvedi *et al.* d'un mode de dissémination original de ces cellules cancéreuses [93].

Par contre, **l'effet du peptide (R/W)<sub>9</sub> sur le phénotype des cellules EF ne semble pas être accompagné de modifications importantes du protéome de ces cellules** : aucune différence significative des niveaux d'expression des protéines n'a pu être mise en évidence dans les conditions d'analyse et de traitement des données utilisées. Une étude du phosphoprotéome ou d'autres modifications post-traductionnelles devrait être envisagée pour comprendre l'impact global sur le fonctionnement des cellules EF. Il serait également utile d'essayer de mieux évaluer l'influence de la durée d'incubation pour la mise en évidence de l'effet du CPP (R/W)<sub>9</sub> sur le protéome.

Pendant de nombreuses années, ce sont les techniques d'études d'interaction protéine-protéine dites *in vitro* (ITC, Cristallographie des rayons X, RMN, spectrométrie de masse native...) qui ont permis de caractériser de nombreux systèmes d'interaction composés de quelques, le plus souvent de deux, partenaires d'interaction purifiés à partir de milieux cellulaires ou obtenus de manière recombinante. Cependant ces techniques deviennent très rapidement inutilisables lorsqu'il s'agit d'étudier des systèmes plus complexes. Dans le cadre de notre étude, nous avons souhaité rechercher sans *a priori* les partenaires d'interaction du peptide vecteur (R/W)<sub>9</sub>, un peptide synthétique sans homologie de séquence avec une molécule existant dans les cellules et n'ayant donc pas de rôle défini. Nous partions ainsi sans réelle piste sur les éventuels partenaires de ce peptide exogène. Etant donnée l'extraordinaire complexité du milieu cellulaire et le peu d'informations disponibles sur le devenir intracellulaire du peptide, cette étude s'est révélée très délicate. Il a été nécessaire de mettre au point différentes approches analytiques complémentaires pour tenter d'obtenir des informations sur le mode d'action du peptide (R/W)<sub>9</sub>. Le point commun de ces approches est l'utilisation de la **spectrométrie de masse** qui constitue aujourd'hui, grâce aux développements instrumentaux et méthodologiques rapides réalisés au cours des dernières décennies, un **outil incontournable pour l'étude des systèmes biologiques**. En effet, le couplage de méthodes de purification d'affinité ou de cross-linking *in cellulo* à la spectrométrie de masse et les approches de protéomique différentielle permettent de travailler sur des systèmes plus complexes et donc de se rapprocher de la « réalité » cellulaire. Cependant, l'obtention d'informations de qualité passe par la maîtrise de plusieurs étapes : conception d'expériences appropriées, préparation d'échantillons, optimisation des conditions d'analyses, traitement des données avec des logiciels adaptés.

Les expériences réalisées ne nous ont pas permis de trancher sur l'origine de l'effet biologique du peptide (R/W)<sub>9</sub> mais ce travail nous a toutefois permis d'évaluer différentes approches et de formuler l'hypothèse d'un impact du peptide (R/W)<sub>9</sub> sur l'activité de transcription de la protéine oncogénique EWS-FLI1. Les résultats obtenus nous incitent également à poursuivre les expériences de protéomique différentielle pour déterminer si (R/W)<sub>9</sub> influe sur les modifications post-traductionnelles des protéines.

J'ai eu la chance avec ce travail de thèse de me familiariser avec différents aspects de la recherche de partenaires d'interactions par AP-MS ou cross-linking-MS et je suis aujourd'hui consciente à la fois du

potentiel immense de ces techniques et de la multiplicité de leurs applications, mais également d'un certain nombre de défis à relever pour faire progresser ce domaine. J'ai pu aussi me rendre compte de **la complexité du milieu cellulaire** et du challenge que constitue l'étude des mécanismes moléculaires induits lors de l'introduction des composés exogènes. Forte de cette expérience, je reste convaincue de la pertinence de ces approches analytiques visant à mieux cerner l'impact de l'entrée de molécules bioactives, CPPs ou médicaments, sur la dynamique cellulaire.



# Matériel et Méthodes

## I. Matériel

### 1. Origine du matériel

Les réactifs utilisés proviennent de Sigma, à part les produits suivants :

- Billes magnétiques de Streptavidine (Dynabeads M-280 Streptavidin, Invitrogen by life Technologies)
- Les gels loaders et les tubes Eppendorfs viennent de Eppendorf
- Consommables pour la culture : Les boîtes de culture (100 X 20 mm) et les pipettes de culture viennent de chez Corning, les boîtes de culture 12 et 96 puits de chez Falcon (Beckton Dickinson Labware).
- Les milieux pour la culture : HBSS, DMEM (high glucose, no pyruvate), le DMEM déplété (high glucose, no glutamine, no lysine, no arginine) pour les expériences de SILAC, les sérums (FBS et NCS), la glutamine (200mM), la pénicilline streptomycine (100 000 UI/L) viennent de Gibco by life technologies.
- Microcons 30 kDa viennent d'Amicon Bioseparations (Millipore).
- Le kit BC Assay vient d'Interchim (Uptima)
- Les acides aminés pour les expériences SILAC (L-Lysine-2HCl; L-Arginine-HCl,  $^{13}\text{C}_6$  L-Lysine-2HCl,  $^{13}\text{C}_6^{15}\text{N}_4$  L-Arginine-HCl) viennent de Thermo Scientific
- L'eau et l'ACN utilisés en NanoLC font partie de la gamme Optima LC/MS et viennent de Fisher Scientific. Le FA et le TFA de Carlo Erba.

### 2. Composition de tampons et solutions couramment utilisés :

#### Tampon de lavages des billes

Tampon A : Tris-HCl 50 mM, 0,1 mg/mL BSA

Tampon B : Tris-HCl 50 mM, 0,1 % SDS, 0,1 mg/mL BSA

Tampon C : Tris-HCl 50mM, 1 M NaCl, 0,1 mg/mL BSA

Bleu de Laemmli 5X : 250 mM Tris, 40 % glycerol, 6,25 % SDS, 0,5 mL mercaptoéthanol 5 %, bleu de bromophénol 0,1 %

Tampon de lyse stringent : (4 % SDS, 100 mM Tris HCl et 100 mM DTT)

#### Solvants utilisés en NanoHPLC

Solvant A : eau contenant 2% d'ACN et 0,1% FA

Solvant B : 98 % d' ACN + 0,1 % FA

## II. Techniques d'analyses et culture cellulaire

### 1. Spectrométrie de masse MALDI-TOF

➤ Préparation de la solution de matrice :

La matrice CHCA est dissoute à 5 mg/mL dans une solution contenant 50 % d'ACN et 0,1 % TFA. La solution obtenue est passée aux ultrasons pendant 5 min afin de faciliter la solubilisation des cristaux de CHCA puis est centrifugée. Les dépôts sur la plaque MALDI sont réalisés par la méthode de la goutte séchée (mélange 1 : 1 échantillon : matrice réalisé au préalable).

➤ Réglages instrumentaux :

Les spectres MALDI-TOF et MALDI-TOF/TOF sont enregistrés sur un spectromètre Applied Biosystems 4700 *Proteomics Analyzer* utilisant un laser NdYAG ( $\lambda = 355$  nm, durée de pulse  $\approx 3$  ns, fréquence de répétition 200Hz). La MS MALDI-TOF est réalisée en mode réflecteur positif (extraction retardée avec une résolution de masse optimisée pour un  $m/z$  2100 u.) et pour des fluences laser proches du seuil de détection des ions. Les spectres sont étalonnés avec une solution d'étalons externes, un mélange de peptides avec des masses comprises entre 500 et 4000 u, déposée à proximité de l'échantillon à analyser.

### 2. Couplage NanoLC-ESI-MS/MS

Le système est un couplage NanoLC (Ultimate 3000, Dionex) avec un spectromètre de masse LTQ-OrbitrapXL (Thermo Scientific) équipé d'une source ETD. Les échantillons sont injectés en mode  $\mu$ L pick-up (boucle d'injection de 20  $\mu$ L  $\rightarrow$  Volume maximum injecté de 6  $\mu$ L) grâce à un passeur d'échantillon et concentrés sur une précolonne (Pepmap,  $C_{18}$ , 300  $\mu$ m  $\times$  5 mm, 5  $\mu$ m, 100 Å, Dionex) avec le solvant A. Après 10 min de chargement, les peptides sont élués sur la colonne NanoLC (Pepmap,  $C_{18}$ , 75  $\mu$ m  $\times$  150 mm, 2  $\mu$ m 100 Å, Dionex) équilibrée avec 98 % de solvant A. Les peptides sont séparés en utilisant un gradient 0-10 min 2-10% de solvant B, 10-70 min 10-25 % de solvant B, 70-90 min 25-40 % de solvant B, 90-93 min 40-60 % de solvant B, 93-110 min 60 % de solvant B et 110-130 min 2% de solvant B à un débit de 300 nL/min. Le spectromètre de masse LTQ-Orbitrap XL est muni d'une source nanoélectrospray (Nano-ESI). L'aiguille d'électrospray de dimensions 360/20  $\mu$ m o.d.  $\times$  10  $\mu$ m i.d. est en silice fondue (PicoTip Emitter, Standard Coated SilicaTip, New Objective). Le spray est obtenu en chauffant le capillaire de transfert à une température de 200 °C et en appliquant une tension typiquement de 1,5 kV entre l'aiguille d'électrospray et le capillaire de transfert. Les spectres MS sont obtenus avec l'analyseur Orbitrap (Automated gain control (AGC)  $2 \times 10^5$ , temps d'injection maximum 500 ms) sur la gamme de  $m/z$  500-2000 avec une résolution de 30k en mode profil, puis des spectres MS/MS sont acquis dans le LTQ (AGC  $1.10^4$ , temps d'injection max 250 ms) en mode DDA (*data dependant acquisition*) en fragmentant les 10 ions les plus intenses avec le mode CID (Top10CID). Un temps d'exclusion dynamique de 180 s est utilisé.

Sauf indication contraire, lorsqu'il est fait mention d'une analyse par NanoLC-ESI-MS/MS cela fait référence à l'utilisation de la méthode décrite ci-dessus.

Pour les milieux de photo-crosslinking, en plus de la méthode classique, des analyses ont été réalisées en utilisant l'outil masstag pour la sélection des précurseurs. De manière à fragmenter

uniquement les pics présents dans les spectres MS sous formes de doublets d'égale intensité et séparés de 10,06 Da, il faut créer dans la rubrique masstag une liste de  $\Delta m/z$  (10,06 ; 5,03 ; 2,52 ; 2,01 ; 1,68) et choisir une gamme de tolérance pour l'intensité du partenaire (70-100 %). L'option « fragmentation des deux partenaires » a été choisie.

Pour la fragmentation avec le mode ETD, une activation supplémentaire (SA) de 20 % a été utilisée.

### 3. Culture cellulaire

Les cellules EF et 3T3 (lignées cellulaires de fibroblastes murins NIH-3T3 respectivement transformées ou non par la protéine de fusion EWS-FLI1) sont des cellules adhérentes cultivées dans du DMEM (+) Glucose (-) Pyruvate complémenté avec 10 % de sérum de veau nouveau-né (NCS), de la pénicilline (100 000 UI/L) et de la streptomycine (100 000 UI/L) sous atmosphère humidifiée contenant 5 % de CO<sub>2</sub>, à 37 °C (incubateur à cellules). Les cultures sont réalisées dans des boîtes de culture de 10 cm de diamètre dans des conditions stériles, toutes les manipulations des cellules se font sous une hotte à flux laminaire. Ces cellulesensemencées ont, dans des conditions de milieu de culture et de dilution classique (entre 1/20<sup>ième</sup> et 1/10<sup>ième</sup>), un temps de doublement d'environ 24 h. Il faut veiller à ne pas trop diluer ces cellules (en particulier les 3T3 qui ne doivent jamais être diluées davantage qu'au 1/20<sup>ième</sup>) et éviter qu'elles ne soient à confluence (en particulier les cellules EF qui sont susceptibles de changer de phénotype à confluence).

## III. Protocoles

Un certain nombre de protocoles ou de partie de protocoles sont communs à différentes expériences. Ceux-ci vont être présentés au début de cette partie de manière à pouvoir ensuite y faire référence dans les protocoles plus spécifiques qui seront, eux, présentés dans l'ordre de leur apparition dans la partie résultats (à l'exception des protocoles relatifs au travail publié (chapitre 4) qui figurent eux directement dans la publication).

### Protocoles communément utilisés :

- Digestion trypsique en solution : L'échantillon à digérer doit être dans un milieu à pH  $\approx$  8 compatible avec la digestion (ex : tampon NH<sub>4</sub>HCO<sub>3</sub> 50 mM). 5  $\mu$ L de solution de DTT à 45 mM dans NH<sub>4</sub>HCO<sub>3</sub> 50 mM est ajoutée et une incubation de 15 min à 50°C est réalisée. Ensuite 5  $\mu$ L de solution d'IAA à 100 mM est ajoutée à l'échantillon qui est incubé 15 min à l'abri de la lumière et à température ambiante. Finalement une solution de trypsine à 0,1  $\mu$ g/ $\mu$ L dans NH<sub>4</sub>HCO<sub>3</sub> 50 mM est ajoutée de manière à avoir un ratio en masse trypsine : protéines de l'ordre de 3 : 100. La digestion est réalisée à 37 °C sur toute une nuit.

- Digestion in-gel : Afin d'éviter toute contamination des échantillons avec de la kératine, il est nécessaire de travailler avec des gants et une charlotte. Les bandes de gel à digérer sont découpées au scalpel sur une surface propre (nettoyage eau puis éthanol) en petits cubes d'environ 1mm<sup>3</sup>. Les morceaux de gels sont placés dans un tube Eppendorf et lavés à deux reprises 15 min sous agitation avec un mélange 50-50 eau/ACN puis une fois 10 min avec de l'ACN de manière à les déshydrater. Une solution de DTT 10 mM dans NH<sub>4</sub>HCO<sub>3</sub> 100 mM est ensuite ajoutée de manière à tout juste recouvrir les morceaux de gels et une incubation de 45 min à 56 °C est réalisée pour réduire les ponts disulfures. Ensuite la solution de DTT est remplacée par une solution d'IAA à 10 mg/mL ( $\approx$ 50 mM)

dans  $\text{NH}_4\text{HCO}_3$  100 mM et une incubation de 30 min à l'abri de la lumière est réalisée. Les morceaux de gels sont ensuite lavés à deux reprises 15 min sous agitation avec un mélange 50-50 eau/ACN puis une fois 10 min avec de l'ACN de manière à les déshydrater. Les tubes sont placés dans la glace et une solution de trypsine à 1,5 ng/ $\mu\text{L}$  dans  $\text{NH}_4\text{HCO}_3$  50 mM est ajoutée. L'incubation dans la glace avec la trypsine dure 45 min afin de laisser le temps à la trypsine de bien diffuser à l'intérieur du gel. L'excédent de solution est éliminé et le gel est recouvert d'une solution de  $\text{NH}_4\text{HCO}_3$  50 mM. Les échantillons sont placés à l'étuve à 37 °C pour une digestion sur la nuit.

A l'issue de la digestion, les échantillons sont centrifugés et les surnageants correspondant aux échantillons digérés sont transférés dans de nouveaux tubes. Si les échantillons doivent être analysés en spectrométrie de masse ou concentrés/ dessalés sur zip tip, ils sont acidifiés de façon à avoir une concentration finale de 0,1% FA. Si les échantillons digérés sont destinés à être purifiés sur billes magnétiques de streptavidine, ils ne doivent pas être acidifiés.

- Conditionnement des billes magnétiques de streptavidine : Les billes magnétiques de streptavidine sont conservées à une concentration de 10 mg/mL dans une solution d'azide de sodium. Elles sont systématiquement conditionnées avant utilisation en réalisant les lavages suivants : 1 fois tampon A, une fois tampon B, une fois tampon C, une fois tampon A avec des volumes environ 4 fois supérieurs au volume de billes à conditionner. Finalement les billes sont reprises dans du tampon A (quatre fois le volume initial) et ajoutées aux milieux réactionnels.

- Lavage classique des billes : A l'issue des incubations des billes avec les milieux réactionnels, celles-ci sont classiquement lavées de la manière suivante : avec deux fois 200  $\mu\text{L}$  de tampon A, deux fois 200  $\mu\text{L}$  de tampon B, deux fois 200  $\mu\text{L}$  de tampon C, puis 200  $\mu\text{L}$ , 100  $\mu\text{L}$  et deux fois 50  $\mu\text{L}$  d'eau. Les billes sont ensuite incubées 5 min avec 50  $\mu\text{L}$  d'une solution de biotine à 10  $\mu\text{M}$  afin de saturer les sites de fixation de la streptavidine et de faciliter l'élution des peptides biotinylés. Elles sont finalement lavées avec 50  $\mu\text{L}$  d'un mélange 1-1 d' $\text{H}_2\text{O}$  et d'ACN.

- Lavage des billes au bleu de Laemmli : Afin d'éliminer un maximum d'interactions non-spécifiques en particulier suite à l'incubation des billes avec des lysats cellulaires non digérés, le protocole classique de lavage des billes a été optimisé : les étapes de lavages avec le tampon B ont été remplacées par une incubation de 10 min à 50 °C avec 50  $\mu\text{L}$  de bleu de Laemmli 1X concentré.

## **1. Evaluation de la cytotoxicité de peptides vecteurs**

La veille de l'expérience (J-1), les cellules sontensemencées en conditions stériles, dans une plaque 96 puits à raison de 20 000 cellules dans 100  $\mu\text{L}$  de milieu de culture par puits. Il faut prévoir 3 puits par concentration de peptides à tester et trois puits pour le contrôle positif (cellules seules). Cette boîte est placée pendant la nuit dans l'incubateur à cellules.

Le jour J, des solutions de peptides aux concentrations à tester sont préparées dans du DMEM à raison de 100  $\mu\text{L}$  de solution peptides pour chaque puits. Lorsque toutes les solutions sont prêtes, le milieu de culture est aspiré. Les cellules sont lavées à deux reprises avec 200  $\mu\text{L}$  de DMEM.

Au temps  $t = 0$  min, les solutions de peptides sont introduites dans les puits. Dans trois puits vides (sans cellules) sont introduits 100  $\mu\text{L}$  de solution de peptide à la concentration la plus forte pour le contrôle négatif. La plaque 96 puits est ensuite replacée dans l'incubateur. Au bout d'une heure, 10

μL de réactif CCK-8 (Dojindo) sont introduits dans chaque puits. Les boîtes 96 puits sont à nouveau placées dans l'incubateur à cellules.

Les mesures d'absorbance sont alors réalisées avec le lecteur de plaques Fluostar à  $t = 2\text{h}$ ,  $3\text{h}30$  et  $5\text{h}$ . Une mesure est réalisée à  $450\text{ nm}$  qui correspond au maximum d'absorbance du composé orangé formazan et une autre mesure est réalisée à  $280\text{ nm}$  correspondant à l'absorbance du plastique de la plaque 96 puits. Les résultats sont récupérés et traités avec le logiciel Excel.

## **2. Evaluation de la quantité de peptide internalisé**

Les expériences d'internalisation sont réalisées dans des plaques 12 puits. La veille,  $5.10^5$  cellules sont ensemencées par puits. Après  $24\text{ h}$  (cellules à confluence), le milieu de culture est aspiré et les cellules sont lavées deux fois avec  $1\text{ mL}$  de DMEM. Les cellules sont ensuite incubées  $75\text{ min}$  avec  $500\text{ }\mu\text{L}$  d'une solution de peptide choisi (photo(R/W)<sub>9</sub> ou (R/W)<sub>9</sub> de référence) à une concentration choisie ( $5$  ou  $7,5\text{ }\mu\text{M}$  dans notre cas) à  $37\text{ }^\circ\text{C}$ .

A la fin de l'incubation, les cellules sont lavées deux fois avec  $1\text{ mL}$  d'HBSS à température ambiante. Les cellules sont décollées et le peptide à la membrane digéré par l'ajout  $500\text{ }\mu\text{L}$  d'une solution de trypsine-EDTA à  $0,05\text{ }\%$  pendant  $5\text{ min}$  à  $37\text{ }^\circ\text{C}$ . La trypsine est inhibée par ajout de  $100\text{ }\mu\text{L}$  d'inhibiteur de trypsine de soja ( $5\text{ mg/mL}$ ) et par  $100\text{ }\mu\text{L}$  de BSA ( $1\text{ mg/mL}$ ). A partir de cette étape, les solutions doivent être stockées et manipulées dans la glace afin d'éviter ou de minimiser toute dégradation.

Le contenu de chaque puits est transféré dans un tube Eppendorf de  $1,5\text{ mL}$ . Chaque puits est ensuite lavé avec  $500\text{ }\mu\text{L}$  de Tris-HCl  $50\text{ mM}$  afin de récupérer les cellules éventuellement restées dans le puits, et cette solution est combinée à la précédente dans le tube Eppendorf correspondant. Les tubes sont centrifugés  $3\text{ min}$  à  $3000\text{ rpm}$  (centrifugeuse Sanyo MSE Micro Centaur) et les surnageants sont éliminés.  $150\text{ }\mu\text{L}$  de solution de lyse pour internalisation ( $0,3\text{ }\%$  Triton X-100, NaCl  $1\text{ M}$ ) contenant la quantité voulue de peptide étalon (= version deutérée du peptide internalisé) sont ajoutés aux culots et les tubes sont immédiatement placés à  $100\text{ }^\circ\text{C}$  pendant  $15\text{ min}$ . Les tubes sont ensuite centrifugés  $5\text{ min}$  à  $13\text{ }000\text{ rpm}$  (centrifugeuse Sanyo MSE Micro Centaur) de façon à culotter les débris membranaires. Les surnageants sont transférés dans des nouveaux tubes Eppendorf de  $1,5\text{ mL}$  et les culots sont lavés avec  $850\text{ }\mu\text{L}$  de Tris-HCl  $50\text{ mM}$ . Cette solution de lavage est combinée au lysat cellulaire correspondant.  $10\text{ }\mu\text{L}$  de suspension de billes magnétiques de streptavidine préalablement conditionnées sont alors ajoutés par tube et une incubation de  $60\text{ min}$  sous agitation douce à température ambiante est réalisée.

Les tubes sont alors placés sur un portoir aimanté permettant d'immobiliser les billes contre la paroi pour les laver. Les billes sont lavées suivant le protocole classique décrit précédemment. Les peptides sont élués des billes magnétiques de streptavidine par l'ajout de  $3\text{ }\mu\text{L}$  de matrice CHCA. Après une incubation de  $10\text{ min}$  en dehors du portoir aimanté et sous agitation, les billes sont à nouveau immobilisées sur l'aimant et  $0,5\text{ }\mu\text{L}$  du surnageant est déposé sur une plaque MALDI avant l'analyse MALDI-TOF.

La quantification du peptide internalisé est réalisée en comparant l'aire du signal MS correspondant au peptide vecteur testé à celle du peptide vecteur étalon utilisé. Les aires des différents pics des massifs isotopiques de chaque signal sont sommées.

### 3. Observation du cytosquelette d'actine des cellules par immunofluorescence

Deux jours avant l'expérience, les lamelles en verre de 12 mm conservées dans de l'éthanol sont réparties dans une boîte 12 puits, lavées une fois avec 500  $\mu$ L de tampon PBS, et recouvertes de 500  $\mu$ L de gélatine à 0,2 %. Cette solution de gélatine à 0,2 % est préparée en dissolvant 0,2 g de gélatine dans 100 mL d'eau (Eau aqua B Brown stérile) dans un bécher qui est ensuite recouvert d'un parafilm et placé quelques minutes aux micro-ondes à la puissance maximale (900 W) afin d'obtenir une solution stérile.

Après 30 min d'incubation, l'essentiel de la solution de gélatine est aspirée et les lamelles sont laissées à sécher sous la hotte toute la nuit.

La veille de l'expérience, 10 000 cellules sont ensemencées en déposant 100  $\mu$ L de solution de cellules formant un ménisque sur la lamelle. Au bout d'une heure, lorsque les cellules ont commencé à adhérer, 1 mL de milieu de culture est ajouté dans chaque puits.

Le jour J, le milieu de culture est aspiré et 500  $\mu$ L de solution de peptides à la concentration à tester (5  $\mu$ M pour photo(R/W)<sub>9</sub>, 10  $\mu$ M pour les peptides (R/W)<sub>9</sub>réf et (R/W)<sub>16</sub>) sont ajoutés pour une incubation de 16 h à 37 °C. A la fin de l'incubation, les cellules sont lavées deux fois avec 500  $\mu$ L de tampon PBS et fixées à l'aide de 500  $\mu$ L d'une solution de paraformaldéhyde à 3 % dans du PBS pendant 30 min à température ambiante. Après deux lavages avec 500  $\mu$ L de tampon PBS, les cellules sont perméabilisées avec 500  $\mu$ L d'une solution de Triton X-100 à 0,4 % pendant 5 min. Les lamelles sont à nouveau lavées deux fois avec du tampon PBS puis incubées 30 minutes avec du sérum de vœu fœtal (FBS) à 10 % dans du PBS afin d'assurer la saturation des sites de fixation non-spécifiques. Le marquage est réalisé sur un morceau de parafilm superposé à un papier absorbant humidifié au PBS pour éviter un assèchement des cellules pendant les différents traitements. Des gouttes de 40  $\mu$ L de phalloïdine FITC 1/100<sup>ième</sup> dans du PBS + 10 % FBS sont déposées sur le parafilm puis les lamelles sont sorties des puits, retournées et délicatement déposées sur ces gouttes avec une pince à épiler. La durée de cette incubation est de 50 min et doit permettre de marquer l'actine-F. Les lamelles sont retournées, les gouttes de phalloïdine aspirées et une incubation de 10 min est réalisée avec 40  $\mu$ L de solution de DAPI au 1/600<sup>ième</sup> dans du PBS+10 % FBS destinée à marquer les noyaux des cellules. On limite ainsi les volumes d'incubation et on maximise le contact entre les lamelles et les composés fluorescents.

Finalement les lamelles sont lavées 2 fois avec du PBS (200  $\mu$ L) avant d'être montées sur lame avec le liquide de montage Vectashield. De petites gouttes de vernis sont disposées tout autour des lamelles de manière à les maintenir en place. Ces lames sont conservées à 4 °C à l'abri de la lumière. Dès le lendemain elles peuvent être observées avec un microscope à fluorescence à filtres (Eclipse, TE 2000-S, Nikon).

### 4. Etude *in vitro* de la réaction de photo-crosslinking

100  $\mu$ L d'une solution à 20  $\mu$ M de BSA et 10  $\mu$ M de chacune de formes, deutérée et non-deutérée, du peptide photo(R/W)<sub>9</sub> dans du NH<sub>4</sub>HCO<sub>3</sub> 50 mM et NaCl 150 mM sont préparés. Cette solution est répartie dans deux tubes Eppendorf. L'un de ces deux tubes est emballé dans du papier aluminium (échantillon contrôle négatif) tandis que l'autre tube est gardé ouvert.

Ces deux tubes sont placés sur un portoir dans un bain eau/glace qui est ensuite introduit dans le montage de photoirradiation. La hauteur de la lampe est réglée de manière à avoir environ 5 cm entre le bas de la lampe et l'échantillon. L'irradiation dure 30 min.

A l'issue de l'irradiation, le milieu réactionnel de photocross-linking ainsi que le milieu contrôle négatif sont réduits, alkylés et digérés à la trypsine toute la nuit à 37 °C.

Des volumes de digests correspondant à 80 pmol d'espèces biotinylées sont repris dans 500 µL de tampon  $\text{NH}_4\text{HCO}_3$  50 mM et incubés avec 20 µL de suspension de billes magnétiques de streptavidine pendant 60 min à température ambiante. Les billes sont ensuite lavées sur un portoir magnétiques en suivant le protocole classique de lavage des billes décrit précédemment. Les peptides biotinylés sont élués avec 5 µL d'HCl 0,1 M. 0,5 µL du surnageant est prélevé, mélangé avec 0,5 µL de matrice CHCA et analysé en MALDI-TOF. Le reste est dilué d'un facteur 2 dans de l'eau et transféré dans un vial pour une analyse NanoLC-ESI-MS/MS.

### 5. Expérience de photocross-linking avec la stratégie N°1

Deux boîtes de cellules EF de 10 cm à confluence (contenant chacune environ  $10^7$  cellules) sont préparées pour l'expérience (ensemencement de  $5 \cdot 10^6$  cellules la veille). Le jour de l'expérience, le milieu de culture est éliminé et remplacé dans une des deux boîtes (échantillon irradié v) par 5 mL d'une solution de peptide photo(R/W)<sub>9</sub> en mélange équimolaire des formes deutérée et non-deutérée dans du DMEM (concentration totale en peptide de 5 µM). Dans l'autre boîte (échantillon contrôle CT), les cellules sont simplement mises en présence de 5 mL de DMEM. Ces boîtes de cellules sont incubées 1h à 37 °C puis placées dans un bain eau/glace dans le montage de photoirradiation à une distance de 5 cm de lampe et irradiées pendant 30 min.

A la fin de l'irradiation, la solution de peptide photo(R/W)<sub>9</sub> ou le DMEM seul sont éliminés et les cellules sont lavées à trois reprises avec 5 mL de PBS de manière à éliminer un maximum de peptide non internalisé. 600 µL de tampon de lyse stringant (4 % SDS, 100 mM Tris HCl et 100 mM DTT) sont ajoutés à chaque boîte de cellules. Les cellules lysées sont décollées du fond des boîtes au moyen d'une raclette et transférées dans des tubes Eppendorf de 1,5 mL qui sont immédiatement placés au bain-marie à 100 °C pour dénaturer les protéases (même si *a priori* les 4 % de SDS sont suffisants pour cela). Les tubes sont ensuite centrifugés 10 min à 16 000 g de manière à culotter les éventuels débris membranaires non solubilisés. Les surnageants sont récupérés et purifiés en plusieurs fois sur des modules de purification par ultracentrifugation (Microcon 30 kDa) en suivant les étapes du protocole FASP (*Filter aided sample preparation*) [324].

#### **Elimination du détergent (*Filter aided sample preparation*) et dosage des protéines :**

Les surnageants des échantillons sont introduits en 3 fois (200 µL par 200 µL) dans la partie supérieure des Microcons complétés à chaque fois par un ajout de 200 µL de tampon UA (8 M urée 100 mM Tris-HCl). Après chaque ajout de surnageant, les modules sont centrifugés 15 min à 14 000 g (le filtrat est régulièrement éliminé). Lorsque la totalité des surnageants a été transférée dans les modules, 100 µL d'une solution IAA (50 mM iodoacétamide, 8 M urée, 100 mM Tris-HCl) sont ajoutés. Les modules sont vortexés 1 min puis incubés 20 min sans agitation à l'abri de la lumière. Après cette étape d'alkylation, les modules sont centrifugés 10 min à 14 000 g. Pour finir, plusieurs



cycles de lavages sont réalisés : 3 cycles de 15 min à 14 000 g après des ajouts de 100  $\mu$ L de tampon UA et 3 cycles de 15 min à 14 000 g après des ajouts de 100  $\mu$ L de  $\text{NH}_4\text{HCO}_3$  100 mM.

A l'issue de ces étapes, les protéines contenues dans la partie supérieure du Microcon sont reprises dans 600  $\mu$ L de tampon  $\text{NH}_4\text{HCO}_3$  50 mM. 100  $\mu$ L seront utilisés pour le dosage des protéines et 250  $\mu$ L sont mis de côté pour la branche a.

Le dosage des protéines du lysat cellulaire par la méthode BCA (BiCinchoninic acid Assay), compatible avec la présence résiduelle de détergents, est réalisé. Typiquement, la lyse de  $10^6$  cellules avec le tampon de lyse stringent permet d'obtenir environ 180  $\mu$ g de protéines.

### **Branche a :**

245 des 250  $\mu$ L d'échantillons v et CT mis de côté à l'issue de la purification sur module d'ultracentrifugation sont mis en présence de 100  $\mu$ L de suspension de billes magnétiques de streptavidine pendant 1 h sous agitation à température ambiante. Les billes sont ensuite lavées sur portoir avec le protocole de lavage des billes au bleu de Laemmli décrit précédemment. Elles sont finalement lavées avec 50  $\mu$ L d'un mélange 1-1 eau-ACN. Puis 30  $\mu$ L de bleu de Laemmli 1X sont ajoutés aux billes qui sont placées au bain marie à 100 °C de manière à éluer les espèces biotinylées. Ces deux échantillons sont déposés sur gel 1D-SDS-PAGE 10 % ainsi que l'équivalent de 10  $\mu$ g de lysats non purifiés sur billes pour chaque condition.

### **Branche b :**

Aux 250  $\mu$ L d'échantillons restants dans les modules d'ultracentrifugation est ajoutée de la trypsine (1:30 en masse). Les modules sont vortexés 1 min puis placés à 37 °C toute la nuit.

Le lendemain, les échantillons digérés sont récupérés dans la partie basse du module d'ultracentrifugation après 10 min de centrifugation à 14 000 g et 3 cycles de lavages avec 50  $\mu$ L de  $\text{NH}_4\text{HCO}_3$  50 mM pendant 10 min à 14 000 g. La moitié de chaque échantillon digéré est purifié sur 50  $\mu$ L de suspension de billes magnétiques de streptavidine (même procédure de lavage des billes que précédemment). Les espèces photocross-linkées sont éluées dans 5  $\mu$ L de HCl 0,1 M pendant 30 min sous agitation. 1  $\mu$ L d'éluat est mélangé à 1  $\mu$ L de matrice CHCA, déposé sur plaque MALDI et analysé en spectrométrie de masse MALDI-TOF.

*Optimisations apportées au protocole afin d'obtenir une digestion enzymatique complète.*

- Après l'irradiation et le lavage des boîtes de cellules au PBS, le tampon de lyse stringent (4 % SDS, 100 mM DTT) est remplacé par le tampon de lyse RIPA (150 mM NaCl, 1% NP-40, 0,5 % déoxycholate de sodium, 0,1 % SDS, 50 mM Tris-HCl).
- Les cellules initialement lysées dans la boîte avec 600  $\mu$ L de détergent, sont décollées mécaniquement à la raclette, transférées dans un tube Eppendorf de 1,5 mL et ensuite seulement lysées par ajout de 150  $\mu$ L de tampon de lyse RIPA.
- Au lieu de la purification sur Microcon, le lysat est repris dans un volume total de 1 mL d'eau puis placé dans un boudin de dialyse (cut-off 5 kDa) et dialysé une nuit contre du tampon  $\text{NH}_4\text{HCO}_3$  50 mM. Cette étape permet de diminuer la concentration en détergents et d'éliminer une partie du peptide photo(R/W)<sub>9</sub> libre, c'est à dire n'ayant pas formé de liaisons covalentes avec des partenaires d'interaction.

## **6. Réalisation d'un western blot pour détecter les espèces biotinylées**

A la fin de la migration des milieux de photocross-linking sur gel 1D SDS-PAGE, ce dernier est électrotransféré sur une membrane de nitrocellulose pendant 2h à ampérage constant de 95mA.

La membrane est ensuite incubée 15 min avec une solution de blocage (PBS/Tween 0,1 %, 1 % BSA) Puis la membrane est mise en présence d'une solution de streptavidine-HRP (ECL Horseradish Peroxydase Streptavidin, Amersham GE Healthcare) au 1/2000<sup>ème</sup> dans du PBS/Tween 0,1 %, 1 % BSA soit 1h à température ambiante, soit toute une nuit à 4 °C, sous agitation douce. La membrane est lavée à 4 ou 5 reprises pendant une dizaine de minutes avec du PBS/Tween 0,1 %.

La membrane est révélée sur un film photographique (Amersham Hyperfilm ECL, GE Healthcare) au moyen d'un kit ECL plus (Kit Perkin Elmer Western lightning ECL plus) et des bains de fixation et de révélation (Sigma Aldrich) en chambre noire.

## **7. Expérience de photocross-linking avec la stratégie N°2**

**Protocole de préfractionnement :**

Tampon 1 : 150 mM NaCl, 50 mM Tris-HCl, 2 mM EDTA

Tampon 2 : 10 mM Tris-HCl, 1 mM EDTA, 1 pastille Complete mini (Roche) pour 15 mL de solution

Tampon 3 : 50 mM Tris-HCl, 1 mM EDTA, 10 mM MgCl<sub>2</sub>, 1 pastille Complete mini (Roche) pour 5 mL de solution, 10 % glycérol

Le préfractionnement cellulaire est typiquement réalisé sur 10 boîtes de cellules de 10 cm à confluence (environ 1.10<sup>8</sup> cellules).

Les cellules sont lavées 2 fois avec 10 mL de tampon 1. 2 mL de tampon 1 sont ajoutés et les cellules sont décollées à l'aide d'une raclette et transférées dans un tube Falcon (5 boîtes de cellules/ tube Falcon). Les tubes Falcons sont centrifugés 15 min à 1500 rpm (Centrifugeuse B3.11 Jouan). Les surnageants sont éliminés. Chaque culot est repris dans 2,5 mL de tampon 2 et transféré dans un potter. Les cellules sont laissées dans la glace 30 min puis sont lysées mécaniquement en réalisant des allers-retours (10-15) au potter. Les cellules lysées sont transférées dans un tube Falcon et centrifugées 10 min à 1000 rpm. Le surnageant est réparti dans deux tubes d'ultracentrifugation. Le culot est ré-extrait à deux reprises avec des volumes décroissants de tampon 2 (2,5 mL puis 1,25 mL) comme précédemment (potterisation et centrifugation) et les surnageants obtenus sont combinés avec les surnageants précédents dans les tubes d'ultracentrifugation.

Le culot restant va constituer la partie « noyau ». Après avoir correctement équilibré les tubes, les surnageants sont soumis à une ultracentrifugation 1 h à 4 °C à 48 000 g (Ultracentrifugeuse Optima L-80 XP). A la fin de l'ultracentrifugation, le surnageant de chaque tube est récupéré dans un tube Falcon et constitue la fraction « cytosol ». Les culots obtenus qui correspondent à la fraction « membranes » sont repris dans un total de 1 mL de tampon 3. Les trois fractions sont dosées avec la méthode BCA. La concentration en protéines de cette fraction « membranes » est d'environ 3 mg/mL. La fraction « membranes » est répartie en 10 aliquotes de 100 µL et conservée à -80°C.

### **Photocross-linking sur membranes :**

3 aliquotes de membranes préalablement préparées sont utilisées (soit un total d'environ 1 mg de protéines membranaires). A deux des trois aliquotes (N°1 et N°2) de membranes sont ajoutés 5 µL d'une solution de peptide photo(R/W)<sub>9</sub> à 100 µM de chaque (concentration finale en peptide photo(R/W)<sub>9</sub> égale à 10 µM). Une de ces aliquotes contenant du peptide (N°1) ainsi que celle n'en contenant pas (N°3) sont photoirradiées 1 h dans un mélange eau/glace. La troisième aliquote (N°2) subit les mêmes traitements mais est emballée dans de l'aluminium pour être protégée du rayonnement UV.

A l'issue de la photoirradiation, les échantillons sont centrifugés 4 min à 13 000 rpm (centrifugeuse Sanyo MSE Micro Centaur). Le surnageant est éliminé et les culots sont repris dans 20 µL de tampon de dénaturation (17 mM DTT, 0,1 % triton X-100, 3 % SDS, 50 mM NH<sub>4</sub>HCO<sub>3</sub>). A ces 20 µL de solution de membranes sont ajoutés 12 µL d'eau mQ et 8 µL de bleu de Laemmli 5X et les échantillons sont ensuite placés 10 min au bain-marie à 100 °C puis répartis entre 2 gels 1D SDS-PAGE 12 %. Après migration, un des deux gels est coloré au bleu de Coomassie et l'autre est transféré sur membrane de nitrocellulose. Un western blot avec de la streptavidine-HRP est ensuite réalisé.

Certaines bandes, dont l'intensité est plus importante pour l'échantillon de membranes photoirradiées en présence de peptide photo(R/W)<sub>9</sub> (N°1) que pour les contrôles (N°2 et 3), sont excisées du gel coloré au bleu. Ces bandes sont digérées à la trypsine et les protéines digérées obtenues sont purifiées sur 10 µL de suspension de billes de streptavidine. Après 60 min d'incubation à température ambiante et sous agitation douce, les billes sont lavées suivant le protocole de lavage au bleu de Laemmli décrit précédemment. Les espèces biotinylées sont éluées avec 5 µL d'HCl 0,1M. 1µL du surnageant est mélangé avec 1µL de matrice CHCA, déposé sur plaque MALDI et analysé en spectrométrie de masse MALDI-TOF.

## **8. Expérience de pull-down + photocross-linking**

Une solution de lyse dite non-dénaturante est fraîchement préparée : 500 µL de NP-40 10 % et une pastille d'inhibiteur de protéase (Complete Mini EDTA free, Roche) sont ajoutés à 4500 µL de PBS 1X. Ainsi la concentration finale en NP-40 dans la solution est de 1 %. Trois boîtes de cellules EF à confluence sont utilisées pour la préparation des lysats cellulaires. Les cellules sont lavées deux fois avec 5 mL d'HBSS puis placées dans la glace et lysées par ajout dans chacune des boîtes de 1 mL de solution de lyse non-dénaturante. Après 1 h d'incubation avec la solution de lyse non-dénaturante, les cellules lysées sont rassemblées avec une raclette et transférées dans des tubes Eppendorf avant d'être centrifugées à 15 000 g (Centrifugeuse HIMAC CT15RE VWR) pendant 10 min à 4°C. Les surnageants obtenus sont rassemblés et dosés par la méthode BCA. Typiquement, la concentration en protéine du lysat obtenu est d'environ 1 mg/mL.

En parallèle de la préparation des lysats cellulaires, 300 µL de billes magnétiques sont conditionnées et réparties de manière égale dans 3 tubes Eppendorf. Aux deux premiers tubes est ajouté 1 mL de solution de peptide photo(R/W)<sub>9</sub> à 0,25 µM de chacune des formes, deutérées et non deutérées, dans NH<sub>4</sub>HCO<sub>3</sub> 50 mM (soit 500 pmol au total de peptide photo(R/W)<sub>9</sub> par tube). Dans le troisième tube est ajouté 1 mL de NH<sub>4</sub>HCO<sub>3</sub> 50 mM. Les tubes sont placés sous agitation à température ambiante pendant 1 h.

A l'issue de l'heure d'incubation, les solutions baignant les billes magnétiques de streptavidine sont éliminées en plaçant les tubes sur portoir aimanté et remplacées à chaque fois par 1 mL de lysat.

Les tubes sont placés sous agitation dans une chambre froide (4 °C) pendant une nuit.

Les échantillons 1 et 3 sont placés au centre de l'arrangement circulaire de 16 néons émettant autour de 350nm du système Rayonet (Southern New England company) et photoirradiés pendant 3 cycles de 20 min. Entre deux cycles les billes qui se sont déposées lors de l'irradiation sont remises en suspension en agitant doucement les tubes.

A la fin des cycles de photoirradiation, les billes sont lavées à 5 reprises avec 500 µL de tampon de lavage doux (25 mM Tris-HCl, 150 mM NaCl, 1 mM EDTA, 1% NP-40, 5 % glycérol) afin d'éliminer une part importante des interactions non-spécifiques établies entre les billes magnétiques de streptavidine greffées ou non par le peptide photo(R/W)<sub>9</sub> et les protéines du lysat cellulaire. Chaque échantillon de billes est ensuite élué avec 40 µL du bleu de Laemmli 2X pendant 10 min à 100 °C. Les éluats sont déposés et migrés sur gel 1D SDS-PAGE 12 %. Le gel est ensuite coloré au bleu de Coomassie.

Chacune des trois pistes correspondant aux trois échantillons est découpée en 10 bandes de 0,5 cm. Les protéines de ces 30 bandes sont digérées in-gel, les échantillons obtenus sont concentrés au speedvac et analysés en NanoLC-ESI-MS/MS.

Les données NanoLC-ESI-MS/MS correspondant aux 10 bandes d'une même piste sont mergées pour la recherche des protéines en banque de données (Banque Uniprot *Mus musculus* au 31/10/13). Les listes de protéines identifiées pour chacune des pistes avec 1 % FDR, sont exportées vers un fichier excel.

Les listes de protéines sont comparées grâce au logiciel SAINT (*Significance Analysis of INteractome*) disponible sur CRAPome.org. Pour la soumission des données via l'interface web il faut construire un fichier au format .csv selon les consignes fournies sur le site.

## 9. Expériences de protéomique différentielle avec la technique SILAC (*Stable Isotope Labelling with Amino acids in cell Culture*)

### ➤ Préparation des milieux de culture :

Les milieux de culture sont à base de DMEM déplété auquel est ajouté du NCS préalablement dialysé (pour éliminer les acides aminés libres) et filtré pour avoir une concentration finale de 10 %, des acides aminés lysine et arginine (146 mg/mL de lysine et 84 mg/mL d'arginine) lourds ou légers selon le milieu à préparer, de la glutamine de manière à avoir une concentration finale de 4mM, une solution de Pénicilline-Streptomycine de manière à avoir une solution 1X concentrée.

Avant d'être introduits dans le milieu de culture les acides aminés sont solubilisés dans 10 mL de DMEM et cette solution est stérilisée par passage sur un filtre de 22 µm.

### ➤ Culture :

Pour les expériences visant à évaluer l'effet d'un traitement avec le peptide photo(R/W)<sub>9</sub> sur les cellules EF, des cellules EF ont été cultivées d'une part avec des acides marqués (H) et d'autre part avec des acides aminés classiques (L).

La culture des cellules pour les deux conditions à comparer doit être réalisée de manière très rigoureuse en s'assurant que les vitesses de croissance des cellules pour les deux conditions sont proches ou, du moins, que les cellules sont à des densités similaires ( $\approx 80\%$  de confluence) le jour de l'expérience.

De manière à avoir une incorporation totale des acides aminés lourds dans les cellules, le début de la culture avec les acides marqués intervient environ 10 jours avant la réalisation de l'expérience (minimum de 6 passages de cellules).

### ➤ Vérification du taux d'incorporation :

Une boîte de chacune des deux conditions à comparer est préparée pour le jour de l'expérience. Les cellules sont lavées avec 10 mL d'HBSS, décollées avec 1 mL de trypsine (5 min dans l'incubateur à cellules) puis la trypsine est stoppée par l'ajout de 5 mL de DMEM déplété contenant 10 % de NCS dialysé. Les cellules sont transférées dans un tube de centrifugation et culottées 5 min à 1000 rpm (Centrifugeuse B3.11 Jouan). Le surnageant est éliminé et les cellules sont reprises dans 4 mL de DMEM déplété. Les cellules sont comptées (Glasstic Slide 10 with grids, Hycor Biomedical) pour chacune des deux conditions. Ensuite les cellules (H) et (L) sont mélangées en ratio 1-1 (au minimum  $10^6$  cellules de chaque). Des échantillons contenant seulement  $10^6$  cellules (L) ou  $10^6$  cellules (H) sont également préparés. Les cellules sont culottées 5 min à 1000 rpm (Centrifugeuse B3.11 Jouan). Le surnageant est éliminé tandis que le culot est repris dans 500  $\mu$ L de PBS 1X et transféré dans un tube Eppendorf. Le culot est lavé à deux reprises avec 1 mL de PBS en centrifugeant entre chaque lavage à 3000 rpm 5 min (Centrifugeuse HIMAC CT 15 RE VWR). Les culots sont ensuite lysés avec un volume de tampon de lyse stringant (4% SDS, 100 mM DTT, 100 mM Tris HCl) choisi de manière à avoir une concentration en protéine dans les lysats d'environ 4  $\mu$ g/ $\mu$ L. Ce volume est calculé sachant qu'en moyenne 1 million de cellules permet d'obtenir 180  $\mu$ g de protéines. Les lysats sont placés au bain marie à 100 °C pendant 15 min puis centrifugés 10 min à 13 000 rpm (Centrifugeuse HIMAC CT 15 RE VWR) afin de culotter d'éventuels débris cellulaires non solubilisés. Les échantillons de lysats pour les cellules (L) seules, (H) seules et pour le mélange 1-1 (H)-(L) sont préparés pour un dépôt sur gel : à 24  $\mu$ L de lysat sont ajoutés 6  $\mu$ L de bleu de Laemmli 5X et l'ensemble est placé au bain-marie à 100 °C pendant 5 min. Les 30  $\mu$ L de chaque échantillon correspondant à 100  $\mu$ g de protéines sont déposés sur un gel 1D SDS-PAGE 12 % de 1,5 mm d'épaisseur. Après migration, le gel est coloré au bleu de Coomassie. Une bande de gel de quelques mm est excisée pour chaque piste, digérée et analysée en spectrométrie de masse MALDI-TOF.

Si le taux d'incorporation est correct ( $> 95\%$ ), la totalité de la piste correspondant au mélange 1-1 (H)-(L) sera découpée, digérée et analysée en NanoLC-ESI-MS/MS de manière à servir d'expérience contrôle.

### ➤ Comparaison des protéomes de cellules EF et 3T3 :

Pour les expériences visant à comparer le protéome des cellules EF avec celui des cellules 3T3, les cellules EF ont été cultivées avec des acides marqués, lourds (H), tandis que les cellules 3T3 ont été cultivées avec des acides aminés classiques, légers (L). Le protocole utilisé est rigoureusement le même que pour la vérification du taux d'incorporation sauf que la piste entière du gel est systématiquement découpée tous les 5 mm et que les 10 ou 11 bandes obtenues sont digérées et analysées en NanoLC-ESI-MS/MS.

### ➤ Test de l'effet d'un peptide vecteur sur le protéome des cellules :

Les cellules lourdes (H) servent de référence. Les cellules légères (L) vont être incubées avec le peptide vecteur à la concentration souhaitée.

*Jour J-1 (pour tester l'effet du peptide avec incubation overnight).* Le peptide vecteur (ex : photo(R/W)<sub>9</sub>) est introduit dans milieu de culture (L) de manière à avoir la concentration finale en peptide souhaitée (ex : 5 µM). La boîte est agitée doucement pendant quelques secondes afin d'avoir une concentration en peptide homogène.

*Jour J :* Le protocole est rigoureusement le même que pour la vérification du taux d'incorporation sauf que la piste entière du gel est systématiquement découpée tous les 5 mm et que les 10 ou 11 bandes obtenues sont digérées pour être analysées en NanoLC-ESI-MS/MS.

### ➤ Traitement des données NanoLC-ESI-MS/MS obtenues :

Les versions L et H des peptides analysés ont été identifiées et quantifiées avec le logiciel Proteome Discoverer 1.3 (Thermo Scientific) avec la carbamidométhylation des cystéines, l'oxydation des méthionines et les acides aminés marqués, lysine <sup>13</sup>C<sub>6</sub> et arginine <sup>13</sup>C<sub>6</sub><sup>15</sup>N<sub>4</sub>, définis comme des modifications variables. Le moteur de recherche SEQUEST a été utilisé avec une banque de données de protéines Uniprot *Mus Musculus* mise à jour le 30/10/13. Le taux de faux positif (FDR) a été fixé à 1 %, la tolérance sur la masse du précurseur à 10 ppm et la tolérance sur la MS/MS à 0,6 Da. Les ratios H/L calculés ont été corrigés avec la normalisation sur le ratio H/L des protéines médianes de façon à prendre en compte l'erreur expérimentale sur le mélange 1-1 des deux conditions.

L'analyse statistique des listes de protéines obtenues pour chacune des expériences a été réalisée avec le logiciel MyProMS [297]. Les valeurs seuils pour le changement du ratio H/L ont été déterminées grâce à une expérience contrôle (mélange 1-1 de deux populations de cellules EF identiques cultivées avec des acides aminés lourds d'une part et légers d'autre part). Il a ainsi été établi que les protéines étaient significativement sur- ou sous-exprimées à partir d'un changement du ratio d'un facteur 1,5 pour des p-values sur le calcul de ces ratios inférieures à 0,05. Pour les expériences visant à comparer (1) les cellules EF et 3T3 et (2) les cellules EF traitées par le peptide (R/W)<sub>9</sub> par rapport aux cellules non traitées, trois réplicats biologiques ont été réalisés.

Seules les protéines présentant une variation de ratio supérieure à un facteur 1,5 en valeur absolue avec une p-value inférieure à 0,05 et quantifiées dans au moins 2 des 3 réplicats ont été conservées pour l'interprétation des données. L'analyse des réseaux d'interactions des protéines significativement sur- ou sous-exprimées a été réalisée avec l'interface web du logiciel STRI





# Références bibliographiques

- [1] J. Howl, S. Matou-Nasri, D. C. West, M. Farquhar, J. Slaninová, C.-G. Ostenson, M. Zorko, P. Ostlund, S. Kumar, U. Langel, J. McKeating, et S. Jones, « Bioportide: an emergent concept of bioactive cell-penetrating peptides », *Cell. Mol. Life Sci. CMLS*, vol. 69, n° 17, p. 2951- 2966, sept. 2012.
- [2] D. Delaroche, F.-X. Cantrelle, F. Subra, C. Van Heijenoort, E. Guittet, C.-Y. Jiao, L. Blanchoin, G. Chassaing, S. Lavielle, C. Auclair, et S. Sagan, « Cell-penetrating peptides with intracellular actin-remodeling activity in malignant fibroblasts », *J. Biol. Chem.*, vol. 285, n° 10, p. 7712- 7721, mars 2010.
- [3] L. H. Hartwell, J. J. Hopfield, S. Leibler, et A. W. Murray, « From molecular to modular cell biology », *Nature*, vol. 402, n° 6761 Suppl, p. C47- 52, déc. 1999.
- [4] C. V. Robinson, A. Sali, et W. Baumeister, « The molecular sociology of the cell », *Nature*, vol. 450, n° 7172, p. 973- 982, déc. 2007.
- [5] S. Bergink et S. Jentsch, « Principles of ubiquitin and SUMO modifications in DNA repair », *Nature*, vol. 458, n° 7237, p. 461- 467, mars 2009.
- [6] R. Ghose, A. Shekhtman, M. J. Goger, H. Ji, et D. Cowburn, « A novel, specific interaction involving the Csk SH3 domain and its natural ligand », *Nat. Struct. Biol.*, vol. 8, n° 11, p. 998- 1004, nov. 2001.
- [7] T. Pawson et J. D. Scott, « Signaling through scaffold, anchoring, and adaptor proteins », *Science*, vol. 278, n° 5346, p. 2075- 2080, déc. 1997.
- [8] T. Pawson et P. Nash, « Assembly of cell regulatory systems through protein interaction domains », *Science*, vol. 300, n° 5618, p. 445- 452, avr. 2003.
- [9] A. Bateman, E. Birney, R. Durbin, S. R. Eddy, K. L. Howe, et E. L. Sonnhammer, « The Pfam protein families database », *Nucleic Acids Res.*, vol. 28, n° 1, p. 263- 266, janv. 2000.
- [10] Y. L. Deribe, T. Pawson, et I. Dikic, « Post-translational modifications in signal integration », *Nat. Struct. Mol. Biol.*, vol. 17, n° 6, p. 666- 672, juin 2010.
- [11] J. Westermarck, J. Ivaska, et G. L. Corthals, « Identification of protein interactions involved in cellular signaling », *Mol. Cell. Proteomics MCP*, vol. 12, n° 7, p. 1752- 1763, juill. 2013.
- [12] T. Obsil et V. Obsilova, « Structural basis of 14-3-3 protein functions », *Semin. Cell Dev. Biol.*, vol. 22, n° 7, p. 663- 672, sept. 2011.
- [13] E. M. Phizicky et S. Fields, « Protein-protein interactions: methods for detection and analysis », *Microbiol. Rev.*, vol. 59, n° 1, p. 94- 123, mars 1995.
- [14] J. Monod, J. Wyman, et J. P. Changeux, « ON THE NATURE OF ALLOSTERIC TRANSITIONS: A PLAUSIBLE MODEL », *J. Mol. Biol.*, vol. 12, p. 88- 118, mai 1965.
- [15] M. Fromont-Racine, J. C. Rain, et P. Legrain, « Toward a functional analysis of the yeast genome through exhaustive two-hybrid screens », *Nat. Genet.*, vol. 16, n° 3, p. 277- 282, juill. 1997.
- [16] A. J. Walhout, R. Sordella, X. Lu, J. L. Hartley, G. F. Temple, M. A. Brasch, N. Thierry-Mieg, et M. Vidal, « Protein interaction mapping in *C. elegans* using proteins involved in vulval development », *Science*, vol. 287, n° 5450, p. 116- 122, janv. 2000.
- [17] H. Goehler, M. Lalowski, U. Stelzl, S. Waelter, M. Stroedicke, U. Worm, A. Droege, K. S. Lindenberg, M. Knoblich, C. Haenig, M. Herbst, J. Suopanki, E. Scherzinger, C. Abraham, B. Bauer, R. Hasenbank, A. Fritzsche, A. H. Ludewig, K. Büssow, K. Buessow, S. H. Coleman, C.-A. Gutekunst, B. G. Landwehrmeyer, H. Lehrach, et E. E. Wanker, « A protein interaction network links GIT1, an enhancer of huntingtin aggregation, to Huntington's disease », *Mol. Cell*, vol. 15, n° 6, p. 853- 865, sept. 2004.

- [18] F. Colland, X. Jacq, V. Trouplin, C. Mougin, C. Groizeleau, A. Hamburger, A. Meil, J. Wojcik, P. Legrain, et J.-M. Gauthier, « Functional proteomics mapping of a human signaling pathway », *Genome Res.*, vol. 14, n° 7, p. 1324- 1332, juill. 2004.
- [19] P. Sheterline, J. Clayton, et J. Sparrow, « Actin », *Protein Profile*, vol. 2, n° 1, p. 1- 103, 1995.
- [20] W. Kabsch, H. G. Mannherz, D. Suck, E. F. Pai, et K. C. Holmes, « Atomic structure of the actin:DNase I complex », *Nature*, vol. 347, n° 6288, p. 37- 44, sept. 1990.
- [21] K. K. Tsuboi, « Actin and bound-nucleotide stoichiometry », *Biochim. Biophys. Acta*, vol. 160, n° 3, p. 420- 434, août 1968.
- [22] C. Frieden et K. Patane, « Differences in G-actin containing bound ATP or ADP: the Mg<sup>2+</sup>-induced conformational change requires ATP », *Biochemistry (Mosc.)*, vol. 24, n° 15, p. 4192- 4196, juill. 1985.
- [23] C. Valentin-Ranc et M. F. Carlier, « Evidence for the direct interaction between tightly bound divalent metal ion and ATP on actin. Binding of the lambda isomers of beta gamma-bidentate CrATP to actin », *J. Biol. Chem.*, vol. 264, n° 35, p. 20871- 20880, déc. 1989.
- [24] H. Lodish, A. Berk, S. L. Zipursky, P. Matsudaira, D. Baltimore, et J. Darnell, « The Actin Cytoskeleton », 2000. [En ligne]. Disponible sur: <http://www.ncbi.nlm.nih.gov/books/NBK21493/>. [Consulté le: 22-juill-2014].
- [25] J. A. Cooper, E. L. Buhle Jr, S. B. Walker, T. Y. Tsong, et T. D. Pollard, « Kinetic evidence for a monomer activation step in actin polymerization », *Biochemistry (Mosc.)*, vol. 22, n° 9, p. 2193- 2202, avr. 1983.
- [26] E. D. Korn, M. F. Carlier, et D. Pantaloni, « Actin polymerization and ATP hydrolysis », *Science*, vol. 238, n° 4827, p. 638- 644, oct. 1987.
- [27] E. M. Bonder, D. J. Fishkind, et M. S. Mooseker, « Direct measurement of critical concentrations and assembly rate constants at the two ends of an actin filament », *Cell*, vol. 34, n° 2, p. 491- 501, sept. 1983.
- [28] M. F. Carlier et D. Pantaloni, « Direct evidence for ADP-Pi-F-actin as the major intermediate in ATP-actin polymerization. Rate of dissociation of Pi from actin filaments », *Biochemistry (Mosc.)*, vol. 25, n° 24, p. 7789- 7792, déc. 1986.
- [29] C. G. dos Remedios, D. Chhabra, M. Kekic, I. V. Dedova, M. Tsubakihara, D. A. Berry, et N. J. Nosworthy, « Actin binding proteins: regulation of cytoskeletal microfilaments », *Physiol. Rev.*, vol. 83, n° 2, p. 433- 473, avr. 2003.
- [30] L. G. Tilney, S. Hatano, H. Ishikawa, et M. S. Mooseker, « The polymerization of actin: its role in the generation of the acrosomal process of certain echinoderm sperm », *J. Cell Biol.*, vol. 59, n° 1, p. 109- 126, oct. 1973.
- [31] D. Safer, M. Elzinga, et V. T. Nachmias, « Thymosin beta 4 and Fx, an actin-sequestering peptide, are indistinguishable », *J. Biol. Chem.*, vol. 266, n° 7, p. 4029- 4032, mars 1991.
- [32] L. Carlsson, L. E. Nyström, I. Sundkvist, F. Markey, et U. Lindberg, « Actin polymerizability is influenced by profilin, a low molecular weight protein in non-muscle cells », *J. Mol. Biol.*, vol. 115, n° 3, p. 465- 483, sept. 1977.
- [33] L. M. Machesky, S. J. Atkinson, C. Ampe, J. Vandekerckhove, et T. D. Pollard, « Purification of a cortical complex containing two unconventional actins from *Acanthamoeba* by affinity chromatography on profilin-agarose », *J. Cell Biol.*, vol. 127, n° 1, p. 107- 115, oct. 1994.
- [34] M. D. Welch, J. Rosenblatt, J. Skoble, D. A. Portnoy, et T. J. Mitchison, « Interaction of human Arp2/3 complex and the *Listeria monocytogenes* ActA protein in actin filament nucleation », *Science*, vol. 281, n° 5373, p. 105- 108, juill. 1998.
- [35] R. D. Mullins, J. A. Heuser, et T. D. Pollard, « The interaction of Arp2/3 complex with actin: nucleation, high affinity pointed end capping, and formation of branching networks of filaments », *Proc. Natl. Acad. Sci. U. S. A.*, vol. 95, n° 11, p. 6181- 6186, mai 1998.
- [36] L. Blanchoin, K. J. Amann, H. N. Higgs, J. B. Marchand, D. A. Kaiser, et T. D. Pollard, « Direct observation of dendritic actin filament networks nucleated by Arp2/3 complex and WASP/Scar proteins », *Nature*, vol. 404, n° 6781, p. 1007- 1011, avr. 2000.

- [37] J. A. Cooper et T. D. Pollard, « Effect of capping protein on the kinetics of actin polymerization », *Biochemistry (Mosc.)*, vol. 24, n° 3, p. 793- 799, janv. 1985.
- [38] H. L. Yin et T. P. Stossel, « Control of cytoplasmic actin gel-sol transformation by gelsolin, a calcium-dependent regulatory protein », *Nature*, vol. 281, n° 5732, p. 583- 586, oct. 1979.
- [39] H. L. Yin, « Gelsolin: calcium- and polyphosphoinositide-regulated actin-modulating protein », *BioEssays News Rev. Mol. Cell. Dev. Biol.*, vol. 7, n° 4, p. 176- 179, oct. 1987.
- [40] M. F. Carlier, F. Ressayre, et D. Pantaloni, « Control of actin dynamics in cell motility. Role of ADF/cofilin », *J. Biol. Chem.*, vol. 274, n° 48, p. 33827- 33830, nov. 1999.
- [41] B. W. Bernstein et J. R. Bamburg, « Tropomyosin binding to F-actin protects the F-actin from disassembly by brain actin-depolymerizing factor (ADF) », *Cell Motil.*, vol. 2, n° 1, p. 1- 8, 1982.
- [42] A. S. McElhinny, S. T. Kazmierski, S. Labeit, et C. C. Gregorio, « Nebulin: the nebulous, multifunctional giant of striated muscle », *Trends Cardiovasc. Med.*, vol. 13, n° 5, p. 195- 201, juill. 2003.
- [43] S. Marston, D. Burton, O. Copeland, I. Fraser, Y. Gao, J. Hodgkinson, P. Huber, B. Levine, M. el-Mezgueldi, et G. Notarianni, « Structural interactions between actin, tropomyosin, caldesmon and calcium binding protein and the regulation of smooth muscle thin filaments », *Acta Physiol. Scand.*, vol. 164, n° 4, p. 401- 414, déc. 1998.
- [44] M. J. McManus, W. L. Lingle, J. L. Salisbury, et N. J. Maihle, « A transformation-associated complex involving tyrosine kinase signal adapter proteins and caldesmon links v-erbB signaling to actin stress fiber disassembly », *Proc. Natl. Acad. Sci. U. S. A.*, vol. 94, n° 21, p. 11351- 11356, oct. 1997.
- [45] C. D. Nobes et A. Hall, « Rho, rac, and cdc42 GTPases regulate the assembly of multimolecular focal complexes associated with actin stress fibers, lamellipodia, and filopodia », *Cell*, vol. 81, n° 1, p. 53- 62, avr. 1995.
- [46] A. van der Flier et A. Sonnenberg, « Structural and functional aspects of filamins », *Biochim. Biophys. Acta*, vol. 1538, n° 2- 3, p. 99- 117, avr. 2001.
- [47] G. H. Thomas, « Spectrin: the ghost in the machine », *BioEssays News Rev. Mol. Cell. Dev. Biol.*, vol. 23, n° 2, p. 152- 160, févr. 2001.
- [48] A. Bretscher et K. Weber, « Villin: the major microfilament-associated protein of the intestinal microvillus », *Proc. Natl. Acad. Sci. U. S. A.*, vol. 76, n° 5, p. 2321- 2325, mai 1979.
- [49] A. Bretscher et K. Weber, « Fimbrin, a new microfilament-associated protein present in microvilli and other cell surface structures », *J. Cell Biol.*, vol. 86, n° 1, p. 335- 340, juill. 1980.
- [50] M. Symons, J. M. Derry, B. Karlak, S. Jiang, V. Lemahieu, F. McCormick, U. Francke, et A. Abo, « Wiskott-Aldrich syndrome protein, a novel effector for the GTPase CDC42Hs, is implicated in actin polymerization », *Cell*, vol. 84, n° 5, p. 723- 734, mars 1996.
- [51] P. Aspenström, U. Lindberg, et A. Hall, « Two GTPases, Cdc42 and Rac, bind directly to a protein implicated in the immunodeficiency disorder Wiskott-Aldrich syndrome », *Curr. Biol. CB*, vol. 6, n° 1, p. 70- 75, janv. 1996.
- [52] L. J. Peterson, Z. Rajfur, A. S. Maddox, C. D. Freel, Y. Chen, M. Edlund, C. Otey, et K. Burridge, « Simultaneous stretching and contraction of stress fibers in vivo », *Mol. Biol. Cell*, vol. 15, n° 7, p. 3497- 3508, juill. 2004.
- [53] M. Rönty, A. Taivainen, M. Moza, C. A. Otey, et O. Carpen, « Molecular analysis of the interaction between palladin and alpha-actinin », *FEBS Lett.*, vol. 566, n° 1- 3, p. 30- 34, mai 2004.
- [54] S. L. Gupton, K. L. Anderson, T. P. Kole, R. S. Fischer, A. Ponti, S. E. Hitchcock-DeGregori, G. Danuser, V. M. Fowler, D. Wirtz, D. Hanein, et C. M. Waterman-Storer, « Cell migration without a lamellipodium: translation of actin dynamics into cell movement mediated by tropomyosin », *J. Cell Biol.*, vol. 168, n° 4, p. 619- 631, févr. 2005.
- [55] P. Naumanen, P. Lappalainen, et P. Hotulainen, « Mechanisms of actin stress fibre assembly », *J. Microsc.*, vol. 231, n° 3, p. 446-454, 2008.
- [56] P. F. Davies, « Flow-mediated endothelial mechanotransduction », *Physiol. Rev.*, vol. 75, n° 3, p. 519- 560, juill. 1995.

- [57] D. R. Carter, G. S. Beaupré, N. J. Giori, et J. A. Helms, « Mechanobiology of skeletal regeneration », *Clin. Orthop.*, n° 355 Suppl, p. S41- 55, oct. 1998.
- [58] D. MacKenna, S. R. Summerour, et F. J. Villarreal, « Role of mechanical factors in modulating cardiac fibroblast function and extracellular matrix synthesis », *Cardiovasc. Res.*, vol. 46, n° 2, p. 257- 263, mai 2000.
- [59] M. Kjaer, « Role of extracellular matrix in adaptation of tendon and skeletal muscle to mechanical loading », *Physiol. Rev.*, vol. 84, n° 2, p. 649- 698, avr. 2004.
- [60] B. Geiger et A. Bershadsky, « Assembly and mechanosensory function of focal contacts », *Curr. Opin. Cell Biol.*, vol. 13, n° 5, p. 584- 592, oct. 2001.
- [61] M. A. Arnaout, S. L. Goodman, et J.-P. Xiong, « Structure and mechanics of integrin-based cell adhesion », *Curr. Opin. Cell Biol.*, vol. 19, n° 5, p. 495- 507, oct. 2007.
- [62] N. Wang, J. P. Butler, et D. E. Ingber, « Mechanotransduction across the cell surface and through the cytoskeleton », *Science*, vol. 260, n° 5111, p. 1124- 1127, mai 1993.
- [63] S. L. Dalton, E. E. Marcantonio, et R. K. Assoian, « Cell attachment controls fibronectin and alpha 5 beta 1 integrin levels in fibroblasts. Implications for anchorage-dependent and - independent growth », *J. Biol. Chem.*, vol. 267, n° 12, p. 8186- 8191, avr. 1992.
- [64] E. Ruoslahti, « Fibronectin and its receptors », *Annu. Rev. Biochem.*, vol. 57, p. 375- 413, 1988.
- [65] D. Choquet, D. P. Felsenfeld, et M. P. Sheetz, « Extracellular matrix rigidity causes strengthening of integrin-cytoskeleton linkages », *Cell*, vol. 88, n° 1, p. 39- 48, janv. 1997.
- [66] S. Miyamoto, S. K. Akiyama, et K. M. Yamada, « Synergistic roles for receptor occupancy and aggregation in integrin transmembrane function », *Science*, vol. 267, n° 5199, p. 883- 885, févr. 1995.
- [67] D. Riveline, E. Zamir, N. Q. Balaban, U. S. Schwarz, T. Ishizaki, S. Narumiya, Z. Kam, B. Geiger, et A. D. Bershadsky, « Focal contacts as mechanosensors: externally applied local mechanical force induces growth of focal contacts by an mDia1-dependent and ROCK-independent mechanism », *J. Cell Biol.*, vol. 153, n° 6, p. 1175- 1186, juin 2001.
- [68] K. Kimura, M. Ito, M. Amano, K. Chihara, Y. Fukata, M. Nakafuku, B. Yamamori, J. Feng, T. Nakano, K. Okawa, A. Iwamatsu, et K. Kaibuchi, « Regulation of myosin phosphatase by Rho and Rho-associated kinase (Rho-kinase) », *Science*, vol. 273, n° 5272, p. 245- 248, juill. 1996.
- [69] D. P. Felsenfeld, P. L. Schwartzberg, A. Venegas, R. Tse, et M. P. Sheetz, « Selective regulation of integrin--cytoskeleton interactions by the tyrosine kinase Src », *Nat. Cell Biol.*, vol. 1, n° 4, p. 200- 206, août 1999.
- [70] M. Glogauer, P. Arora, G. Yao, I. Sokholov, J. Ferrier, et C. A. McCulloch, « Calcium ions and tyrosine phosphorylation interact coordinately with actin to regulate cytoprotective responses to stretching », *J. Cell Sci.*, vol. 110 ( Pt 1), p. 11- 21, janv. 1997.
- [71] M. Sokabe, K. Naruse, S. Sai, T. Yamada, K. Kawakami, M. Inoue, K. Murase, et M. Miyazu, « Mechanotransduction and intracellular signaling mechanisms of stretch-induced remodeling in endothelial cells », *Heart Vessels*, vol. Suppl 12, p. 191- 193, 1997.
- [72] G. Pawlak et D. M. Helfman, « Cytoskeletal changes in cell transformation and tumorigenesis », *Curr. Opin. Genet. Dev.*, vol. 11, n° 1, p. 41- 47, févr. 2001.
- [73] M. J. Renan, « How many mutations are required for tumorigenesis? Implications from human cancer data », *Mol. Carcinog.*, vol. 7, n° 3, p. 139- 146, 1993.
- [74] D. Hanahan et R. A. Weinberg, « The hallmarks of cancer », *Cell*, vol. 100, n° 1, p. 57- 70, janv. 2000.
- [75] D. Hanahan et R. A. Weinberg, « Hallmarks of cancer: the next generation », *Cell*, vol. 144, n° 5, p. 646- 674, mars 2011.
- [76] P. K. Mattila et P. Lappalainen, « Filopodia: molecular architecture and cellular functions », *Nat. Rev. Mol. Cell Biol.*, vol. 9, n° 6, p. 446- 454, juin 2008.
- [77] A. M. Weaver, « Invadopodia: specialized cell structures for cancer invasion », *Clin. Exp. Metastasis*, vol. 23, n° 2, p. 97- 105, 2006.
- [78] S. R. Gross, « Actin binding proteins: their ups and downs in metastatic life », *Cell Adhes. Migr.*, vol. 7, n° 2, p. 199- 213, avr. 2013.

- [79] J. Ewing, « Classics in oncology. Diffuse endothelioma of bone. James Ewing. Proceedings of the New York Pathological Society, 1921 », *CA. Cancer J. Clin.*, vol. 22, n° 2, p. 95- 98, avr. 1972.
- [80] C. A. Arndt et W. M. Crist, « Common musculoskeletal tumors of childhood and adolescence », *N. Engl. J. Med.*, vol. 341, n° 5, p. 342- 352, juill. 1999.
- [81] H. E. Grier, « The Ewing family of tumors. Ewing's sarcoma and primitive neuroectodermal tumors », *Pediatr. Clin. North Am.*, vol. 44, n° 4, p. 991- 1004, août 1997.
- [82] M. E. Kadin et K. G. Bensh, « On the origin of Ewing's tumor », *Cancer*, vol. 27, n° 2, p. 257- 273, févr. 1971.
- [83] P. S. Dickman, L. A. Liotta, et T. J. Triche, « Ewing's sarcoma. Characterization in established cultures and evidence of its histogenesis », *Lab. Invest. J. Tech. Methods Pathol.*, vol. 47, n° 4, p. 375- 382, oct. 1982.
- [84] J. J. Navas-Palacios, R. Aparicio-Duque, et M. D. Valdés, « On the histogenesis of Ewing's sarcoma. An ultrastructural, immunohistochemical, and cytochemical study », *Cancer*, vol. 53, n° 9, p. 1882- 1901, mai 1984.
- [85] M. S. Staeger, C. Hutter, I. Neumann, S. Foja, U. E. Hattenhorst, G. Hansen, D. Afar, et S. E. G. Burdach, « DNA microarrays reveal relationship of Ewing family tumors to both endothelial and fetal neural crest-derived cells and define novel targets », *Cancer Res.*, vol. 64, n° 22, p. 8213- 8221, nov. 2004.
- [86] O. Delattre, J. Zucman, B. Plougastel, C. Desmaze, T. Melot, M. Peter, H. Kovar, I. Joubert, P. de Jong, et G. Rouleau, « Gene fusion with an ETS DNA-binding domain caused by chromosome translocation in human tumours », *Nature*, vol. 359, n° 6391, p. 162- 165, sept. 1992.
- [87] M. Bernstein, H. Kovar, M. Paulussen, R. L. Randall, A. Schuck, L. A. Teot, et H. Juergens, « Ewing's sarcoma family of tumors: current management », *The Oncologist*, vol. 11, n° 5, p. 503- 519, mai 2006.
- [88] A. Arvand et C. T. Denny, « Biology of EWS/ETS fusions in Ewing's family tumors », *Oncogene*, vol. 20, n° 40, p. 5747- 5754, sept. 2001.
- [89] A. Prieur, F. Tirode, P. Cohen, et O. Delattre, « EWS/FLI-1 silencing and gene profiling of Ewing cells reveal downstream oncogenic pathways and a crucial role for repression of insulin-like growth factor binding protein 3 », *Mol. Cell. Biol.*, vol. 24, n° 16, p. 7275- 7283, août 2004.
- [90] A. Uren, O. Tcherkasskaya, et J. A. Toretsky, « Recombinant EWS-FLI1 oncoprotein activates transcription », *Biochemistry (Mosc.)*, vol. 43, n° 42, p. 13579- 13589, oct. 2004.
- [91] K. P. Ng, G. Potikyan, R. O. V. Savene, C. T. Denny, V. N. Uversky, et K. A. W. Lee, « Multiple aromatic side chains within a disordered structure are critical for transcription and transforming activity of EWS family oncoproteins », *Proc. Natl. Acad. Sci. U. S. A.*, vol. 104, n° 2, p. 479- 484, janv. 2007.
- [92] H. V. Erkizan, V. N. Uversky, et J. A. Toretsky, « Oncogenic partnerships: EWS-FLI1 protein interactions initiate key pathways of Ewing's sarcoma », *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.*, vol. 16, n° 16, p. 4077- 4083, août 2010.
- [93] A. Chaturvedi, L. M. Hoffman, A. L. Welm, S. L. Lessnick, et M. C. Beckerle, « The EWS/FLI Oncogene Drives Changes in Cellular Morphology, Adhesion, and Migration in Ewing Sarcoma », *Genes Cancer*, vol. 3, n° 2, p. 102- 116, févr. 2012.
- [94] V. Amsellem, M.-H. Kryszke, M. Hervy, F. Subra, R. Athman, H. Leh, C. Brachet-Ducos, et C. Auclair, « The actin cytoskeleton-associated protein zyxin acts as a tumor suppressor in Ewing tumor cells », *Exp. Cell Res.*, vol. 304, n° 2, p. 443- 456, avr. 2005.
- [95] Y. Song, R. S. Maul, C. S. Gerbin, et D. D. Chang, « Inhibition of anchorage-independent growth of transformed NIH3T3 cells by epithelial protein lost in neoplasm (EPLIN) requires localization of EPLIN to actin cytoskeleton », *Mol. Biol. Cell*, vol. 13, n° 4, p. 1408- 1416, avr. 2002.
- [96] P. Braun et A.-C. Gingras, « History of protein-protein interactions: from egg-white to complex networks », *Proteomics*, vol. 12, n° 10, p. 1478- 1498, mai 2012.



- [97] A. Grigoriev, « On the number of protein-protein interactions in the yeast proteome », *Nucleic Acids Res.*, vol. 31, n° 14, p. 4157- 4161, juill. 2003.
- [98] S. Fields et O. Song, « A novel genetic system to detect protein-protein interactions », *Nature*, vol. 340, n° 6230, p. 245- 246, juill. 1989.
- [99] C. T. Chien, P. L. Bartel, R. Sternglanz, et S. Fields, « The two-hybrid system: a method to identify and clone genes for proteins that interact with a protein of interest », *Proc. Natl. Acad. Sci. U. S. A.*, vol. 88, n° 21, p. 9578- 9582, nov. 1991.
- [100] S. Cabantous, T. C. Terwilliger, et G. S. Waldo, « Protein tagging and detection with engineered self-assembling fragments of green fluorescent protein », *Nat. Biotechnol.*, vol. 23, n° 1, p. 102- 107, janv. 2005.
- [101] T. K. Kerppola, « Visualization of molecular interactions by fluorescence complementation », *Nat. Rev. Mol. Cell Biol.*, vol. 7, n° 6, p. 449- 456, juin 2006.
- [102] U. Stelzl, U. Worm, M. Lalowski, C. Haenig, F. H. Brembeck, H. Goehler, M. Stroedicke, M. Zenkner, A. Schoenherr, S. Koeppen, J. Timm, S. Mintzlaff, C. Abraham, N. Bock, S. Kietzmann, A. Goedde, E. Toksöz, A. Droege, S. Krobitsch, B. Korn, W. Birchmeier, H. Lehrach, et E. E. Wanker, « A human protein-protein interaction network: a resource for annotating the proteome », *Cell*, vol. 122, n° 6, p. 957- 968, sept. 2005.
- [103] B. Liedberg, C. Nylander, et I. Lundström, « Biosensing with surface plasmon resonance--how it all started », *Biosens. Bioelectron.*, vol. 10, n° 8, p. i- ix, 1995.
- [104] M. Malmqvist, « Biospecific interaction analysis using biosensor technology », *Nature*, vol. 361, n° 6408, p. 186- 187, janv. 1993.
- [105] U. Jönsson, L. Fägerstam, B. Ivarsson, B. Johnsson, R. Karlsson, K. Lundh, S. Löfås, B. Persson, H. Roos, et I. Rönnerberg, « Real-time biospecific interaction analysis using surface plasmon resonance and a sensor chip technology », *BioTechniques*, vol. 11, n° 5, p. 620- 627, nov. 1991.
- [106] A. Karlsson-Borgå et W. Rolfsen, « Methodological considerations when using nitrocellulose immunoblotting from polyacrylamide gels to study the mould allergens *Aspergillus fumigatus* and *Alternaria alternata* », *J. Immunol. Methods*, vol. 136, n° 1, p. 91- 102, janv. 1991.
- [107] G. Panayotou, G. Gish, P. End, O. Truong, I. Gout, R. Dhand, M. J. Fry, I. Hiles, T. Pawson, et M. D. Waterfield, « Interactions between SH2 domains and tyrosine-phosphorylated platelet-derived growth factor beta-receptor sequences: analysis of kinetic parameters by a novel biosensor-based approach », *Mol. Cell. Biol.*, vol. 13, n° 6, p. 3567- 3576, juin 1993.
- [108] J. Kikuchi, Y. Furukawa, et N. Hayashi, « Identification of novel p53-binding proteins by biomolecular interaction analysis combined with tandem mass spectrometry », *Mol. Biotechnol.*, vol. 23, n° 3, p. 203- 212, mars 2003.
- [109] H. Schägger et G. von Jagow, « Blue native electrophoresis for isolation of membrane protein complexes in enzymatically active form », *Anal. Biochem.*, vol. 199, n° 2, p. 223- 231, déc. 1991.
- [110] H. Eubel, L. Jänsch, et H.-P. Braun, « New insights into the respiratory chain of plant mitochondria. Supercomplexes and a unique composition of complex II », *Plant Physiol.*, vol. 133, n° 1, p. 274- 286, sept. 2003.
- [111] I. Arnold, K. Pfeiffer, W. Neupert, R. A. Stuart, et H. Schägger, « Yeast mitochondrial F1F0-ATP synthase exists as a dimer: identification of three dimer-specific subunits », *EMBO J.*, vol. 17, n° 24, p. 7170- 7178, déc. 1998.
- [112] I. Wittig, H.-P. Braun, et H. Schägger, « Blue native PAGE : Article : Nature Protocols », *Nat Protoc.*, vol. 1, n° 1, p. 418- 428, juin 2006.
- [113] Y. V. Miteva, H. G. Budayeva, et I. M. Cristea, « Proteomics-based methods for discovery, quantification, and validation of protein-protein interactions », *Anal. Chem.*, vol. 85, n° 2, p. 749- 768, janv. 2013.
- [114] B. Brizzard, « Epitope tagging », *BioTechniques*, vol. 44, n° 5, p. 693- 695, avr. 2008.

- [115] L. Trinkle-Mulcahy, « Resolving protein interactions and complexes by affinity purification followed by label-based quantitative mass spectrometry », *Proteomics*, vol. 12, n° 10, p. 1623- 1638, mai 2012.
- [116] N. J. Moorman, R. Sharon-Friling, T. Shenk, et I. M. Cristea, « A targeted spatial-temporal proteomics approach implicates multiple cellular trafficking pathways in human cytomegalovirus virion maturation », *Mol. Cell. Proteomics MCP*, vol. 9, n° 5, p. 851- 860, mai 2010.
- [117] G. Rigaut, A. Shevchenko, B. Rutz, M. Wilm, M. Mann, et B. Séraphin, « A generic protein purification method for protein complex characterization and proteome exploration », *Nat. Biotechnol.*, vol. 17, n° 10, p. 1030- 1032, oct. 1999.
- [118] M. O. Collins et J. S. Choudhary, « Mapping multiprotein complexes by affinity purification and mass spectrometry », *Curr. Opin. Biotechnol.*, vol. 19, n° 4, p. 324- 330, août 2008.
- [119] A. C. Gingras, M. Gstaiger, B. Raught, et R. Aebersold, « Analysis of protein complexes using mass spectrometry », *Nat. Rev. Mol. Cell Biol.*, vol. 8, n° 8, p. 645-654, 2007.
- [120] G. Drewes, « Chemical proteomics in drug discovery », *Methods Mol. Biol. Clifton NJ*, vol. 803, p. 15- 21, 2012.
- [121] M. P. Washburn, D. Wolters, et J. R. Yates, « Large-scale analysis of the yeast proteome by multidimensional protein identification technology », *Nat. Biotechnol.*, vol. 19, n° 3, p. 242- 247, mars 2001.
- [122] J. K. Eng, A. L. McCormack, et J. R. Yates, « An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database », *J. Am. Soc. Mass Spectrom.*, vol. 5, n° 11, p. 976- 989, nov. 1994.
- [123] D. N. Perkins, D. J. Pappin, D. M. Creasy, et J. S. Cottrell, « Probability-based protein identification by searching sequence databases using mass spectrometry data », *Electrophoresis*, vol. 20, n° 18, p. 3551- 3567, déc. 1999.
- [124] A. D. Catherman, O. S. Skinner, et N. L. Kelleher, « Top Down proteomics: facts and perspectives », *Biochem. Biophys. Res. Commun.*, vol. 445, n° 4, p. 683- 693, mars 2014.
- [125] L. M. Smith, N. L. Kelleher, et Consortium for Top Down Proteomics, « Proteoform: a single term describing protein complexity », *Nat. Methods*, vol. 10, n° 3, p. 186- 187, mars 2013.
- [126] J. C. Tran, L. Zamdborg, D. R. Ahlf, J. E. Lee, A. D. Catherman, K. R. Durbin, J. D. Tipton, A. Vellaichamy, J. F. Kellie, M. Li, C. Wu, S. M. M. Sweet, B. P. Early, N. Siuti, R. D. LeDuc, P. D. Compton, P. M. Thomas, et N. L. Kelleher, « Mapping intact protein isoforms in discovery mode using top-down proteomics », *Nature*, vol. 480, n° 7376, p. 254- 258, déc. 2011.
- [127] Z. Wu, A. Moghaddas Gholami, et B. Kuster, « Systematic identification of the HSP90 candidate regulated proteome », *Mol. Cell. Proteomics MCP*, vol. 11, n° 6, p. M111.016675, juin 2012.
- [128] I. M. Cristea, R. Williams, B. T. Chait, et M. P. Rout, « Fluorescent proteins as proteomic probes », *Mol. Cell. Proteomics MCP*, vol. 4, n° 12, p. 1933- 1941, déc. 2005.
- [129] S. Jäger, N. Gulbahce, P. Cimermancic, J. Kane, N. He, S. Chou, I. D'Orso, J. Fernandes, G. Jang, A. D. Frankel, T. Alber, Q. Zhou, et N. J. Krogan, « Purification and characterization of HIV-human protein complexes », *Methods San Diego Calif*, vol. 53, n° 1, p. 13- 19, janv. 2011.
- [130] M. Brajenovic, G. Joberty, B. Küster, T. Bouwmeester, et G. Drewes, « Comprehensive proteomic analysis of human Par protein complexes reveals an interconnected protein network », *J. Biol. Chem.*, vol. 279, n° 13, p. 12804- 12811, mars 2004.
- [131] D. V. Skarra, M. Goudreault, H. Choi, M. Mullin, A. I. Nesvizhskii, A.-C. Gingras, et R. E. Honkanen, « Label-free quantitative proteomics and SAINT analysis enable interactome mapping for the human Ser/Thr protein phosphatase 5 », *Proteomics*, vol. 11, n° 8, p. 1508- 1516, avr. 2011.
- [132] A.-C. Gavin, P. Aloy, P. Grandi, R. Krause, M. Boesche, M. Marzioch, C. Rau, L. J. Jensen, S. Bastuck, B. Dimpelfeld, A. Edelmann, M.-A. Heurtier, V. Hoffman, C. Hoefert, K. Klein, M. Hudak, A.-M. Michon, M. Schelder, M. Schirle, M. Remor, T. Rudi, S. Hooper, A. Bauer, T. Bouwmeester, G. Casari, G. Drewes, G. Neubauer, J. M. Rick, B. Kuster, P. Bork, R. B. Russell, et



- G. Superti-Furga, « Proteome survey reveals modularity of the yeast cell machinery », *Nature*, vol. 440, n° 7084, p. 631- 636, mars 2006.
- [133] N. J. Krogan, G. Cagney, H. Yu, G. Zhong, X. Guo, A. Ignatchenko, J. Li, S. Pu, N. Datta, A. P. Tikuisis, T. Punna, J. M. Peregrín-Alvarez, M. Shales, X. Zhang, M. Davey, M. D. Robinson, A. Paccanaro, J. E. Bray, A. Sheung, B. Beattie, D. P. Richards, V. Canadien, A. Lalev, F. Mena, P. Wong, A. Starostine, M. M. Canete, J. Vlasblom, S. Wu, C. Orsi, S. R. Collins, S. Chandran, R. Haw, J. J. Rilstone, K. Gandi, N. J. Thompson, G. Musso, P. St Onge, S. Ghanny, M. H. Y. Lam, G. Butland, A. M. Altaf-Ul, S. Kanaya, A. Shilatifard, E. O'Shea, J. S. Weissman, C. J. Ingles, T. R. Hughes, J. Parkinson, M. Gerstein, S. J. Wodak, A. Emili, et J. F. Greenblatt, « Global landscape of protein complexes in the yeast *Saccharomyces cerevisiae* », *Nature*, vol. 440, n° 7084, p. 637- 643, mars 2006.
- [134] M. Pardo et J. S. Choudhary, « Assignment of protein interactions from affinity purification/mass spectrometry data », *J. Proteome Res.*, vol. 11, n° 3, p. 1462- 1474, mars 2012.
- [135] H. Choi, B. Larsen, Z.-Y. Lin, A. Breitskreutz, D. Mellacheruvu, D. Fermin, Z. S. Qin, M. Tyers, A.-C. Gingras, et A. I. Nesvizhskii, « SAINT: probabilistic scoring of affinity purification-mass spectrometry data », *Nat. Methods*, vol. 8, n° 1, p. 70- 73, janv. 2011.
- [136] A. Breitskreutz, H. Choi, J. R. Sharom, L. Boucher, V. Neduva, B. Larsen, Z.-Y. Lin, B.-J. Breitskreutz, C. Stark, G. Liu, J. Ahn, D. Dewar-Darch, T. Regul, X. Tang, R. Almeida, Z. S. Qin, T. Pawson, A.-C. Gingras, A. I. Nesvizhskii, et M. Tyers, « A global protein kinase and phosphatase interaction network in yeast », *Science*, vol. 328, n° 5981, p. 1043- 1046, mai 2010.
- [137] S. Mädler, C. Bich, D. Touboul, et R. Zenobi, « Chemical cross-linking with NHS esters: a systematic study on amino acid reactivities », *J. Mass Spectrom. JMS*, vol. 44, n° 5, p. 694- 706, mai 2009.
- [138] S. M. Chowdhury, X. Du, N. Tolić, S. Wu, R. J. Moore, M. U. Mayer, R. D. Smith, et J. N. Adkins, « Identification of cross-linked peptides after click-based enrichment using sequential collision-induced dissociation and electron transfer dissociation tandem mass spectrometry », *Anal. Chem.*, vol. 81, n° 13, p. 5524- 5532, juill. 2009.
- [139] C. H. Sohn, H. D. Agnew, J. E. Lee, M. J. Sweredoski, R. L. J. Graham, G. T. Smith, S. Hess, G. Czerwieniec, J. A. Loo, J. R. Heath, R. J. Deshaies, et J. L. Beauchamp, « Designer reagents for mass spectrometry-based proteomics: clickable cross-linkers for elucidation of protein structures and interactions », *Anal. Chem.*, vol. 84, n° 6, p. 2662- 2669, mars 2012.
- [140] X. Tang, G. R. Munske, W. F. Siems, et J. E. Bruce, « Mass spectrometry identifiable cross-linking strategy for studying protein-protein interactions », *Anal. Chem.*, vol. 77, n° 1, p. 311- 318, janv. 2005.
- [141] M. Q. Müller, J. J. Zeiser, F. Dreiocker, A. Pich, M. Schäfer, et A. Sinz, « A universal matrix-assisted laser desorption/ionization cleavable cross-linker for protein structure analysis », *Rapid Commun. Mass Spectrom. RCM*, vol. 25, n° 1, p. 155- 161, janv. 2011.
- [142] C. Ihling, A. Schmidt, S. Kalkhof, D. M. Schulz, C. Stingl, K. Mechtler, M. Haack, A. G. Beck-Sickinger, D. M. F. Cooper, et A. Sinz, « Isotope-labeled cross-linkers and Fourier transform ion cyclotron resonance mass spectrometry for structural analysis of a protein/peptide complex », *J. Am. Soc. Mass Spectrom.*, vol. 17, n° 8, p. 1100- 1113, août 2006.
- [143] D. Paramelle, S. Cantel, C. Enjalbal, M. Amblard, E. Forest, M. Heymann, C. Geourjon, J. Martinez, et G. Subra, « A new generation of cross-linkers for selective detection by MALDI MS », *Proteomics*, vol. 9, n° 23, p. 5384- 5388, déc. 2009.
- [144] E. Sachon, G. Bolbach, G. Chassaing, S. Lavielle, et S. Sagan, « Cgamma H2 of Met174 side chain is the site of covalent attachment of a substance P analog photoactivable in position 5 », *J. Biol. Chem.*, vol. 277, n° 52, p. 50409- 50414, déc. 2002.
- [145] O. Rinner, J. Seebacher, T. Walzthoeni, L. N. Mueller, M. Beck, A. Schmidt, M. Mueller, et R. Aebersold, « Identification of cross-linked peptides from large sequence databases », *Nat. Methods*, vol. 5, n° 4, p. 315- 318, avr. 2008.

- [146] R. Fritzsche, C. H. Ihling, M. Götze, et A. Sinz, « Optimizing the enrichment of cross-linked products for mass spectrometric protein analysis », *Rapid Commun. Mass Spectrom. RCM*, vol. 26, n° 6, p. 653- 658, mars 2012.
- [147] A. Leitner, T. Walzthoeni, A. Kahraman, F. Herzog, O. Rinner, M. Beck, et R. Aebersold, « Probing Native Protein Structures by Chemical Cross-Linking, Mass Spectrometry, and Bioinformatics », *Mol. Cell. Proteomics*, vol. 9, n° 8, p. 1634- 1649, janv. 2010.
- [148] C. A. G. Söderberg, W. Lambert, S. Kjellström, A. Wiegandt, R. P. Wulff, C. Månsson, G. Rutsdottir, et C. Emanuelsson, « Detection of Crosslinks within and between Proteins by LC-MALDI-TOF/TOF and the Software FINDX to Reduce the MS/MS-Data to Acquire for Validation », *PLoS One*, vol. 7, n° 6, p. e38927, 2012.
- [149] M. I. Rasmussen, J. C. Refsgaard, L. Peng, G. Houen, et P. Højrup, « CrossWork: software-assisted identification of cross-linked peptides », *J. Proteomics*, vol. 74, n° 10, p. 1871- 1883, sept. 2011.
- [150] X. Du, S. M. Chowdhury, N. P. Manes, S. Wu, M. U. Mayer, J. N. Adkins, G. A. Anderson, et R. D. Smith, « Xlink-identifier: an automated data analysis platform for confident identifications of chemically cross-linked peptides using tandem mass spectrometry », *J. Proteome Res.*, vol. 10, n° 3, p. 923- 931, mars 2011.
- [151] W. Lambert, G. Rutsdottir, R. Hussein, K. Bernfur, S. Kjellström, et C. Emanuelsson, « Probing the transient interaction between the small heat-shock protein Hsp21 and a model substrate protein using crosslinking mass spectrometry », *Cell Stress Chaperones*, vol. 18, n° 1, p. 75- 85, janv. 2013.
- [152] B. X. Huang et H.-Y. Kim, « Effective identification of Akt interacting proteins by two-step chemical crosslinking, co-immunoprecipitation and mass spectrometry », *PLoS One*, vol. 8, n° 4, p. e61430, 2013.
- [153] H. Pertl-Obermeyer, W. X. Schulze, et G. Obermeyer, « In vivo cross-linking combined with mass spectrometry analysis reveals receptor-like kinases and Ca(2+) signalling proteins as putative interaction partners of pollen plasma membrane H(+) ATPases », *J. Proteomics*, vol. 108, p. 17- 29, août 2014.
- [154] B. W. Sutherland, J. Toews, et J. Kast, « Utility of formaldehyde cross-linking and mass spectrometry in the study of protein-protein interactions », *J. Mass Spectrom. JMS*, vol. 43, n° 6, p. 699- 715, juin 2008.
- [155] J. W. Chin et P. G. Schultz, « In vivo photocrosslinking with unnatural amino acid mutagenesis », *ChemBiochem Eur. J. Chem. Biol.*, vol. 3, n° 11, p. 1135- 1137, nov. 2002.
- [156] M. Suchanek, A. Radzikowska, et C. Thiele, « Photo-leucine and photo-methionine allow identification of protein-protein interactions in living cells », *Nat. Methods*, vol. 2, n° 4, p. 261- 267, avr. 2005.
- [157] J. J. Christensen, R. M. Izatt, et D. Eatough, « Thermodynamics of Metal Cyanide Coordination. V. Log K,  $\Delta H^\circ$ , and  $\Delta S^\circ$  Values for the  $Hg_2^{2+}$ -CN<sup>-</sup> System », *Inorg. Chem.*, vol. 4, n° 9, p. 1278- 1280, sept. 1965.
- [158] M. W. Freyer et E. A. Lewis, « Isothermal titration calorimetry: experimental design, data analysis, and probing macromolecule/ligand binding and kinetic interactions », *Methods Cell Biol.*, vol. 84, p. 79- 113, 2008.
- [159] M. J. Cliff, A. Gutierrez, et J. E. Ladbury, « A survey of the year 2003 literature on applications of isothermal titration calorimetry », *J. Mol. Recognit. JMR*, vol. 17, n° 6, p. 513- 523, déc. 2004.
- [160] Y. Liang, « Applications of isothermal titration calorimetry in protein science », *Acta Biochim. Biophys. Sin.*, vol. 40, n° 7, p. 565- 576, juill. 2008.
- [161] M. R. Arkin et J. A. Wells, « Small-molecule inhibitors of protein-protein interactions: progressing towards the dream », *Nat. Rev. Drug Discov.*, vol. 3, n° 4, p. 301- 317, avr. 2004.
- [162] C. K. Liew, K. Kowalski, A. H. Fox, A. Newton, B. K. Sharpe, M. Crossley, et J. P. Mackay, « Solution structures of two CCHC zinc fingers from the FOG family protein U-shaped that

- mediate protein-protein interactions », *Struct. Lond. Engl.* 1993, vol. 8, n° 11, p. 1157- 1166, nov. 2000.
- [163] P. Mercier, M. X. Li, et B. D. Sykes, « Role of the structural domain of troponin C in muscle regulation: NMR studies of Ca<sup>2+</sup> binding and subsequent interactions with regions 1-40 and 96-115 of troponin I », *Biochemistry (Mosc.)*, vol. 39, n° 11, p. 2902- 2911, mars 2000.
- [164] I. R. Kleckner et M. P. Foster, « An introduction to NMR-based approaches for measuring protein dynamics », *Biochim. Biophys. Acta*, vol. 1814, n° 8, p. 942- 968, août 2011.
- [165] A. M. Davis, S. A. St-Gallay, et G. J. Kleywegt, « Limitations and lessons in the use of X-ray structural information in drug design », *Drug Discov. Today*, vol. 13, n° 19- 20, p. 831- 841, oct. 2008.
- [166] J.-P. Renaud et M.-A. Delsuc, « Biophysical techniques for ligand screening and drug design », *Curr. Opin. Pharmacol.*, vol. 9, n° 5, p. 622- 628, oct. 2009.
- [167] G. Bolbach, « Matrix-assisted laser desorption/ionization analysis of non-covalent complexes: fundamentals and applications », *Curr. Pharm. Des.*, vol. 11, n° 20, p. 2535- 2557, 2005.
- [168] B. Ganem, Y. T. Li, et J. D. Henion, « Observation of noncovalent enzyme-substrate and enzyme-product complexes by ion-spray mass spectrometry », *J. Am. Chem. Soc.*, vol. 113, n° 20, p. 7818-7819, 1991.
- [169] V. Katta et B. T. Chait, « Observation of the heme-globin complex in native myoglobin by electrospray-ionization mass spectrometry », *J. Am. Chem. Soc.*, vol. 113, n° 22, p. 8534-8535, 1991.
- [170] S. Sanglier, C. Atmanene, G. Chevreux, et A. V. Dorsselaer, « Nondenaturing mass spectrometry to study noncovalent protein/protein and protein/ligand complexes: technical aspects and application to the determination of binding stoichiometries », *Methods Mol. Biol. Clifton NJ*, vol. 484, p. 217- 243, 2008.
- [171] A. Ayed, A. N. Krutchinsky, W. Ens, K. G. Standing, et H. W. Duckworth, « Quantitative evaluation of protein-protein and ligand-protein equilibria of a large allosteric enzyme by electrospray ionization time-of-flight mass spectrometry », *Rapid Commun. Mass Spectrom. RCM*, vol. 12, n° 7, p. 339- 344, 1998.
- [172] M. C. Fitzgerald, I. Chernushevich, K. G. Standing, C. P. Whitman, et S. B. Kent, « Probing the oligomeric structure of an enzyme by electrospray ionization time-of-flight mass spectrometry », *Proc. Natl. Acad. Sci. U. S. A.*, vol. 93, n° 14, p. 6851- 6856, juill. 1996.
- [173] A. J. R. Heck, « Native mass spectrometry: a bridge between interactomics and structural biology », *Nat. Methods*, vol. 5, n° 11, p. 927- 933, nov. 2008.
- [174] V. Katta et B. T. Chait, « Conformational changes in proteins probed by hydrogen-exchange electrospray-ionization mass spectrometry », *Rapid Commun. Mass Spectrom. RCM*, vol. 5, n° 4, p. 214- 217, avr. 1991.
- [175] I. A. Kaltashov, C. E. Bobst, et R. R. Abzalimov, « H/D exchange and mass spectrometry in the studies of protein conformation and dynamics: is there a need for a top-down approach? », *Anal. Chem.*, vol. 81, n° 19, p. 7892- 7899, oct. 2009.
- [176] J. Pan, J. Han, C. H. Borchers, et L. Konermann, « Characterizing short-lived protein folding intermediates by top-down hydrogen exchange mass spectrometry », *Anal. Chem.*, vol. 82, n° 20, p. 8591- 8597, oct. 2010.
- [177] T. D. Tullius et B. A. Dombroski, « Hydroxyl radical "footprinting": high-resolution information about DNA-protein contacts and application to lambda repressor and Cro protein », *Proc. Natl. Acad. Sci. U. S. A.*, vol. 83, n° 15, p. 5469- 5473, août 1986.
- [178] J. G. Kiselar et M. R. Chance, « Future directions of structural mass spectrometry using hydroxyl radical footprinting », *J. Mass Spectrom. JMS*, vol. 45, n° 12, p. 1373- 1382, déc. 2010.
- [179] M. Bantscheff, M. Schirle, G. Sweetman, J. Rick, et B. Kuster, « Quantitative mass spectrometry in proteomics: a critical review », *Anal. Bioanal. Chem.*, vol. 389, n° 4, p. 1017- 1031, oct. 2007.

- [180] S.-E. Ong et M. Mann, « Mass spectrometry-based proteomics turns quantitative », *Nat. Chem. Biol.*, vol. 1, n° 5, p. 252- 262, oct. 2005.
- [181] T. Geiger, J. Cox, P. Ostasiewicz, J. R. Wisniewski, et M. Mann, « Super-SILAC mix for quantitative proteomics of human tumor tissue », *Nat. Methods*, vol. 7, n° 5, p. 383- 385, mai 2010.
- [182] T. Geiger, J. R. Wisniewski, J. Cox, S. Zanivan, M. Kruger, Y. Ishihama, et M. Mann, « Use of stable isotope labeling by amino acids in cell culture as a spike-in standard in quantitative proteomics », *Nat. Protoc.*, vol. 6, n° 2, p. 147- 157, févr. 2011.
- [183] M. Unlü, M. E. Morgan, et J. S. Minden, « Difference gel electrophoresis: a single gel method for detecting changes in protein extracts », *Electrophoresis*, vol. 18, n° 11, p. 2071- 2077, oct. 1997.
- [184] A. Thompson, J. Schäfer, K. Kuhn, S. Kienle, J. Schwarz, G. Schmidt, T. Neumann, R. Johnstone, A. K. A. Mohammed, et C. Hamon, « Tandem mass tags: a novel quantification strategy for comparative analysis of complex protein mixtures by MS/MS », *Anal. Chem.*, vol. 75, n° 8, p. 1895- 1904, avr. 2003.
- [185] S. Wiese, K. A. Reidegeld, H. E. Meyer, et B. Warscheid, « Protein labeling by iTRAQ: a new tool for quantitative mass spectrometry in proteome research », *Proteomics*, vol. 7, n° 3, p. 340- 350, févr. 2007.
- [186] L. V. DeSouza, A. M. Taylor, W. Li, M. S. Minkoff, A. D. Romaschin, T. J. Colgan, et K. W. M. Siu, « Multiple reaction monitoring of mTRAQ-labeled peptides enables absolute quantification of endogenous levels of a potential cancer marker in cancerous and normal endometrial tissues », *J. Proteome Res.*, vol. 7, n° 8, p. 3525- 3534, août 2008.
- [187] P. Mertins, N. D. Udeshi, K. R. Clauser, D. R. Mani, J. Patel, S. Ong, J. D. Jaffe, et S. A. Carr, « iTRAQ labeling is superior to mTRAQ for quantitative global proteomics and phosphoproteomics », *Mol. Cell. Proteomics MCP*, vol. 11, n° 6, p. M111.014423, juin 2012.
- [188] J.-L. Hsu, S.-Y. Huang, N.-H. Chow, et S.-H. Chen, « Stable-isotope dimethyl labeling for quantitative proteomics », *Anal. Chem.*, vol. 75, n° 24, p. 6843- 6852, déc. 2003.
- [189] P. J. Boersema, T. T. Aye, T. A. B. van Veen, A. J. R. Heck, et S. Mohammed, « Triplex protein quantification based on stable isotope labeling by peptide dimethylation applied to cell and tissue lysates », *Proteomics*, vol. 8, n° 22, p. 4624- 4632, nov. 2008.
- [190] P. V. Bondarenko, D. Chelius, et T. A. Shaler, « Identification and relative quantitation of protein mixtures by enzymatic digestion followed by capillary reversed-phase liquid chromatography-tandem mass spectrometry », *Anal. Chem.*, vol. 74, n° 18, p. 4741- 4749, sept. 2002.
- [191] R. Costenoble, P. Picotti, L. Reiter, R. Stallmach, M. Heinemann, U. Sauer, et R. Aebersold, « Comprehensive quantitative analysis of central carbon and amino-acid metabolism in *Saccharomyces cerevisiae* under multiple conditions by targeted proteomics », *Mol. Syst. Biol.*, vol. 7, p. 464, févr. 2011.
- [192] J. V. Olsen, M. Vermeulen, A. Santamaria, C. Kumar, M. L. Miller, L. J. Jensen, F. Gnäd, J. Cox, T. S. Jensen, E. A. Nigg, S. Brunak, et M. Mann, « Quantitative phosphoproteomics reveals widespread full phosphorylation site occupancy during mitosis », *Sci. Signal.*, vol. 3, n° 104, p. ra3, 2010.
- [193] D. W. Huang, B. T. Sherman, et R. A. Lempicki, « Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources », *Nat. Protoc.*, vol. 4, n° 1, p. 44- 57, 2009.
- [194] D. W. Huang, B. T. Sherman, et R. A. Lempicki, « Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists », *Nucleic Acids Res.*, vol. 37, n° 1, p. 1- 13, janv. 2009.
- [195] A. Franceschini, D. Szklarczyk, S. Frankild, M. Kuhn, M. Simonovic, A. Roth, J. Lin, P. Minguez, P. Bork, C. von Mering, et L. J. Jensen, « STRING v9.1: protein-protein interaction networks, with increased coverage and integration », *Nucleic Acids Res.*, vol. 41, n° Database issue, p. D808- 815, janv. 2013.

- [196] D. Szklarczyk, A. Franceschini, M. Kuhn, M. Simonovic, A. Roth, P. Minguéz, T. Doerks, M. Stark, J. Muller, P. Bork, L. J. Jensen, et C. von Mering, « The STRING database in 2011: functional interaction networks of proteins, globally integrated and scored », *Nucleic Acids Res.*, vol. 39, n° Database issue, p. D561- D568, janv. 2011.
- [197] A. Bensimon, A. J. R. Heck, et R. Aebersold, « Mass spectrometry-based proteomics and network biology », *Annu. Rev. Biochem.*, vol. 81, p. 379- 405, 2012.
- [198] L. C. Gillet, P. Navarro, S. Tate, H. Rost, N. Selevsek, L. Reiter, R. Bonner, et R. Aebersold, « Targeted Data Extraction of the MS/MS Spectra Generated by Data-independent Acquisition: A New Concept for Consistent and Accurate Proteome Analysis », *Mol. Cell. Proteomics MCP*, vol. 11, n° 6, juin 2012.
- [199] B. C. Collins, L. C. Gillet, G. Rosenberger, H. L. Röst, A. Vichalkovski, M. Gstaiger, et R. Aebersold, « Quantifying protein interaction dynamics by SWATH mass spectrometry: application to the 14-3-3 system », *Nat. Methods*, vol. 10, n° 12, p. 1246- 1253, déc. 2013.
- [200] J.-P. Lambert, G. Ivosev, A. L. Couzens, B. Larsen, M. Taipale, Z.-Y. Lin, Q. Zhong, S. Lindquist, M. Vidal, R. Aebersold, T. Pawson, R. Bonner, S. Tate, et A.-C. Gingras, « Mapping differential interactomes by affinity purification coupled with data-independent mass spectrometry acquisition », *Nat. Methods*, vol. 10, n° 12, p. 1239- 1245, déc. 2013.
- [201] A. D. Frankel et C. O. Pabo, « Cellular uptake of the tat protein from human immunodeficiency virus », *Cell*, vol. 55, n° 6, p. 1189- 1193, déc. 1988.
- [202] M. Green et P. M. Loewenstein, « Autonomous functional domains of chemically synthesized human immunodeficiency virus tat trans-activator protein », *Cell*, vol. 55, n° 6, p. 1179- 1188, déc. 1988.
- [203] E. Vivès, P. Brodin, et B. Lebleu, « A truncated HIV-1 Tat protein basic domain rapidly translocates through the plasma membrane and accumulates in the cell nucleus », *J. Biol. Chem.*, vol. 272, n° 25, p. 16010- 16017, juin 1997.
- [204] A. Joliot, C. Pernelle, H. Deagostini-Bazin, et A. Prochiantz, « Antennapedia homeobox peptide regulates neural morphogenesis », *Proc. Natl. Acad. Sci. U. S. A.*, vol. 88, n° 5, p. 1864- 1868, mars 1991.
- [205] D. Derossi, A. H. Joliot, G. Chassaing, et A. Prochiantz, « The third helix of the Antennapedia homeodomain translocates through biological membranes. », *J. Biol. Chem.*, vol. 269, n° 14, p. 10444- 10450, août 1994.
- [206] J. Oehlke, M. Beyermann, B. Wiesner, M. Melzig, H. Berger, E. Krause, et M. Bienert, « Evidence for extensive and non-specific translocation of oligopeptides across plasma membranes of mammalian cells », *Biochim. Biophys. Acta*, vol. 1330, n° 1, p. 50- 60, nov. 1997.
- [207] C. Bechara et S. Sagan, « Cell-penetrating peptides: 20 years later, where do we stand? », *FEBS Lett.*, vol. 587, n° 12, p. 1693- 1702, juin 2013.
- [208] A. Elmquist, M. Lindgren, T. Bartfai, et U. Langel, « VE-cadherin-derived cell-penetrating peptide, pVEC, with carrier functions », *Exp. Cell Res.*, vol. 269, n° 2, p. 237- 244, oct. 2001.
- [209] M. Pooga, M. Lindgren, M. Hällbrink, E. Bråkenhielm, et U. Langel, « Galanin-based peptides, galparan and transportan, with receptor-dependent and independent activities », *Ann. N. Y. Acad. Sci.*, vol. 863, p. 450- 453, déc. 1998.
- [210] M. C. Morris, P. Vidal, L. Chaloin, F. Heitz, et G. Divita, « A new peptide vector for efficient delivery of oligonucleotides into mammalian cells », *Nucleic Acids Res.*, vol. 25, n° 14, p. 2730- 2736, juill. 1997.
- [211] M. C. Morris, J. Depollier, J. Mery, F. Heitz, et G. Divita, « A peptide carrier for the delivery of biologically active proteins into mammalian cells », *Nat. Biotechnol.*, vol. 19, n° 12, p. 1173- 1176, déc. 2001.
- [212] D. J. Mitchell, D. T. Kim, L. Steinman, C. G. Fathman, et J. B. Rothbard, « Polyarginine enters cells more efficiently than other polycationic homopolymers », *J. Pept. Res. Off. J. Am. Pept. Soc.*, vol. 56, n° 5, p. 318- 325, nov. 2000.



- [213] S. Futaki, T. Suzuki, W. Ohashi, T. Yagami, S. Tanaka, K. Ueda, et Y. Sugiura, « Arginine-rich peptides. An abundant source of membrane-permeable peptides having potential as carriers for intracellular protein delivery », *J. Biol. Chem.*, vol. 276, n° 8, p. 5836- 5840, févr. 2001.
- [214] J. Oehlke, A. Scheller, B. Wiesner, E. Krause, M. Beyermann, E. Klauschenz, M. Melzig, et M. Bienert, « Cellular uptake of an alpha-helical amphipathic model peptide with the potential to deliver polar compounds into the cell interior non-endocytically », *Biochim. Biophys. Acta*, vol. 1414, n° 1- 2, p. 127- 139, nov. 1998.
- [215] D. Delaroche, B. Aussedat, S. Aubry, G. Chassaing, F. Burlina, G. Clodic, G. Bolbach, S. Lavielle, et S. Sagan, « Tracking a new cell-penetrating (W/R) nonapeptide, through an enzyme-stable mass spectrometry reporter tag », *Anal. Chem.*, vol. 79, n° 5, p. 1932- 1938, mars 2007.
- [216] D. Derossi, S. Calvet, A. Trembleau, A. Brunissen, G. Chassaing, et A. Prochiantz, « Cell internalization of the third helix of the Antennapedia homeodomain is receptor-independent », *J. Biol. Chem.*, vol. 271, n° 30, p. 18188- 18193, juill. 1996.
- [217] J. P. Richard, K. Melikov, E. Vives, C. Ramos, B. Verbeure, M. J. Gait, L. V. Chernomordik, et B. Lebleu, « Cell-penetrating peptides. A reevaluation of the mechanism of cellular uptake », *J. Biol. Chem.*, vol. 278, n° 1, p. 585- 590, janv. 2003.
- [218] M. Fotin-Mleczek, R. Fischer, et R. Brock, « Endocytosis and cationic cell-penetrating peptides-a merger of concepts and methods », *Curr. Pharm. Des.*, vol. 11, n° 28, p. 3613- 3628, 2005.
- [219] I. M. Kaplan, J. S. Wadia, et S. F. Dowdy, « Cationic TAT peptide transduction domain enters cells by macropinocytosis », *J. Control. Release Off. J. Control. Release Soc.*, vol. 102, n° 1, p. 247- 253, janv. 2005.
- [220] I. Nakase, M. Niwa, T. Takeuchi, K. Sonomura, N. Kawabata, Y. Koike, M. Takehashi, S. Tanaka, K. Ueda, J. C. Simpson, A. T. Jones, Y. Sugiura, et S. Futaki, « Cellular uptake of arginine-rich peptides: roles for macropinocytosis and actin rearrangement », *Mol. Ther. J. Am. Soc. Gene Ther.*, vol. 10, n° 6, p. 1011- 1022, déc. 2004.
- [221] H. Binder et G. Lindblom, « Charge-dependent translocation of the Trojan peptide penetratin across lipid membranes », *Biophys. J.*, vol. 85, n° 2, p. 982- 995, août 2003.
- [222] D. Terrone, S. L. W. Sang, L. Roudaia, et J. R. Silvius, « Penetratin and related cell-penetrating cationic peptides can translocate across lipid bilayers in the presence of a transbilayer potential », *Biochemistry (Mosc.)*, vol. 42, n° 47, p. 13787- 13799, déc. 2003.
- [223] F. Madani, S. Lindberg, U. Langel, S. Futaki, et A. Gräslund, « Mechanisms of cellular uptake of cell-penetrating peptides », *J. Biophys. Hindawi Publ. Corp. Online*, vol. 2011, p. 414729, 2011.
- [224] C.-Y. Jiao, D. Delaroche, F. Burlina, I. D. Alves, G. Chassaing, et S. Sagan, « Translocation and Endocytosis for Cell-Penetrating Peptide Internalization », *J. Biol. Chem.*, vol. 284, n° 49, p. 33957- 33965, avr. 2009.
- [225] J. B. Rothbard, T. C. Jessop, R. S. Lewis, B. A. Murray, et P. A. Wender, « Role of membrane potential and hydrogen bonding in the mechanism of translocation of guanidinium-rich peptides into cells », *J. Am. Chem. Soc.*, vol. 126, n° 31, p. 9506- 9507, août 2004.
- [226] I. Nakase, T. Takeuchi, G. Tanaka, et S. Futaki, « Methodological and cellular aspects that govern the internalization mechanisms of arginine-rich cell-penetrating peptides », *Adv. Drug Deliv. Rev.*, vol. 60, n° 4- 5, p. 598- 607, mars 2008.
- [227] C. N. Carrigan et B. Imperiali, « The engineering of membrane-permeable peptides », *Anal. Biochem.*, vol. 341, n° 2, p. 290- 298, juin 2005.
- [228] A. Elmquist, M. Hansen, et U. Langel, « Structure-activity relationship study of the cell-penetrating peptide pVEC », *Biochim. Biophys. Acta*, vol. 1758, n° 6, p. 721- 729, juin 2006.
- [229] C. Bechara, M. Pallerla, Y. Zaltsman, F. Burlina, I. D. Alves, O. Lequin, et S. Sagan, « Tryptophan within basic peptide sequences triggers glycosaminoglycan-dependent endocytosis », *FASEB J. Off. Publ. Fed. Am. Soc. Exp. Biol.*, vol. 27, n° 2, p. 738- 749, févr. 2013.
- [230] A. Prochiantz, « Messenger proteins: homeoproteins, TAT and others », *Curr. Opin. Cell Biol.*, vol. 12, n° 4, p. 400- 406, août 2000.
- [231] W. P. R. Verdurmen, P. H. Bovee-Geurts, P. Wadhwani, A. S. Ulrich, M. Hällbrink, T. H. van Kuppevelt, et R. Brock, « Preferential uptake of L- versus D-amino acid cell-penetrating

- peptides in a cell type-dependent manner », *Chem. Biol.*, vol. 18, n° 8, p. 1000- 1010, août 2011.
- [232] E. Gonçalves, E. Kitas, et J. Seelig, « Structural and thermodynamic aspects of the interaction between heparan sulfate and analogues of melittin », *Biochemistry (Mosc.)*, vol. 45, n° 9, p. 3086- 3094, mars 2006.
- [233] H. C. Christianson et M. Belting, « Heparan sulfate proteoglycan as a cell-surface endocytosis receptor », *Matrix Biol. J. Int. Soc. Matrix Biol.*, vol. 35, p. 51- 55, avr. 2014.
- [234] M. Tyagi, M. Rusnati, M. Presta, et M. Giacca, « Internalization of HIV-1 tat requires cell surface heparan sulfate proteoglycans », *J. Biol. Chem.*, vol. 276, n° 5, p. 3254- 3261, févr. 2001.
- [235] M. Belting, « Heparan sulfate proteoglycan as a plasma membrane carrier », *Trends Biochem. Sci.*, vol. 28, n° 3, p. 145- 151, mars 2003.
- [236] M. Yanagishita, « Metabolism of plasma membrane-associated heparan sulfate proteoglycans », *Adv. Exp. Med. Biol.*, vol. 313, p. 113- 120, 1992.
- [237] A. Ziegler et J. Seelig, « Binding and clustering of glycosaminoglycans: a common property of mono- and multivalent cell-penetrating compounds », *Biophys. J.*, vol. 94, n° 6, p. 2142- 2149, mars 2008.
- [238] R. G. Parton et K. Simons, « The multiple faces of caveolae », *Nat. Rev. Mol. Cell Biol.*, vol. 8, n° 3, p. 185- 194, mars 2007.
- [239] M. C. Kerr et R. D. Teasdale, « Defining macropinocytosis », *Traffic Cph. Den.*, vol. 10, n° 4, p. 364- 371, avr. 2009.
- [240] S. A. Mousavi, L. Malerød, T. Berg, et R. Kjekshus, « Clathrin-dependent endocytosis », *Biochem. J.*, vol. 377, n° Pt 1, p. 1- 16, janv. 2004.
- [241] J. P. Richard, K. Melikov, H. Brooks, P. Prevot, B. Lebleu, et L. V. Chernomordik, « Cellular uptake of unconjugated TAT peptide involves clathrin-dependent endocytosis and heparan sulfate receptors », *J. Biol. Chem.*, vol. 280, n° 15, p. 15300- 15306, avr. 2005.
- [242] A. Ferrari, V. Pellegrini, C. Arcangeli, A. Fittipaldi, M. Giacca, et F. Beltram, « Caveolae-mediated internalization of extracellular HIV-1 tat fusion proteins visualized in real time », *Mol. Ther. J. Am. Soc. Gene Ther.*, vol. 8, n° 2, p. 284- 294, août 2003.
- [243] A. Walrant, I. Correia, C.-Y. Jiao, O. Lequin, E. H. Bent, N. Goasdoué, C. Lacombe, G. Chassaing, S. Sagan, et I. D. Alves, « Different membrane behaviour and cellular uptake of three basic arginine-rich peptides », *Biochim. Biophys. Acta BBA - Biomembr.*, vol. 1808, n° 1, p. 382- 393, janv. 2011.
- [244] P. Säälk, A. Elmquist, M. Hansen, K. Padari, K. Saar, K. Viht, U. Langel, et M. Pooga, « Protein cargo delivery properties of cell-penetrating peptides. A comparative study », *Bioconjug. Chem.*, vol. 15, n° 6, p. 1246- 1253, déc. 2004.
- [245] I. D. Alves, N. Goasdoué, I. Correia, S. Aubry, C. Galanth, S. Sagan, S. Lavielle, et G. Chassaing, « Membrane interaction and perturbation mechanisms induced by two cationic cell penetrating peptides with distinct charge distribution », *Biochim. Biophys. Acta*, vol. 1780, n° 7- 8, p. 948- 959, août 2008.
- [246] P. Joanne, C. Galanth, N. Goasdoué, P. Nicolas, S. Sagan, S. Lavielle, G. Chassaing, C. El Amri, et I. D. Alves, « Lipid reorganization induced by membrane-active peptides probed using differential scanning calorimetry », *Biochim. Biophys. Acta*, vol. 1788, n° 9, p. 1772- 1781, sept. 2009.
- [247] H. D. Herce, A. E. Garcia, J. Litt, R. S. Kane, P. Martin, N. Enrique, A. Rebolledo, et V. Milesi, « Arginine-rich peptides destabilize the plasma membrane, consistent with a pore formation translocation mechanism of cell-penetrating peptides », *Biophys. J.*, vol. 97, n° 7, p. 1917- 1925, oct. 2009.
- [248] I. Nakase, H. Hirose, G. Tanaka, A. Tadokoro, S. Kobayashi, T. Takeuchi, et S. Futaki, « Cell-surface Accumulation of Flock House Virus-derived Peptide Leads to Efficient Internalization via Macropinocytosis », *Mol. Ther.*, vol. 17, n° 11, p. 1868- 1876, août 2009.



- [249] H. Hirose, T. Takeuchi, H. Osakada, S. Pujals, S. Katayama, I. Nakase, S. Kobayashi, T. Haraguchi, et S. Futaki, « Transient focal membrane deformation induced by arginine-rich peptides leads to their direct penetration into cells », *Mol. Ther. J. Am. Soc. Gene Ther.*, vol. 20, n° 5, p. 984- 993, mai 2012.
- [250] A. Mishra, G. H. Lai, N. W. Schmidt, V. Z. Sun, A. R. Rodriguez, R. Tong, L. Tang, J. Cheng, T. J. Deming, D. T. Kamei, et G. C. L. Wong, « Translocation of HIV TAT peptide and analogues induced by multiplexed membrane and cytoskeletal interactions », *Proc. Natl. Acad. Sci.*, vol. 108, n° 41, p. 16883- 16888, nov. 2011.
- [251] H. H. Szeto, P. W. Schiller, K. Zhao, et G. Luo, « Fluorescent dyes alter intracellular targeting and function of cell-penetrating tetrapeptides », *FASEB J. Off. Publ. Fed. Am. Soc. Exp. Biol.*, vol. 19, n° 1, p. 118- 120, janv. 2005.
- [252] C. A. Puckett et J. K. Barton, « Fluorescein redirects a ruthenium-octaarginine conjugate to the nucleus », *J. Am. Chem. Soc.*, vol. 131, n° 25, p. 8738- 8739, juill. 2009.
- [253] J.-R. Bertrand, C. Malvy, T. Auguste, G. K. Tóth, O. Kiss-Ivánkóvits, E. Illyés, M. Hollósi, S. Bottka, et I. Laczkó, « Synthesis and studies on cell-penetrating peptides », *Bioconjug. Chem.*, vol. 20, n° 7, p. 1307- 1314, juill. 2009.
- [254] A. Ziegler et J. Seelig, « High affinity of the cell-penetrating peptide HIV-1 Tat-PTD for DNA », *Biochemistry (Mosc.)*, vol. 46, n° 27, p. 8138- 8145, juill. 2007.
- [255] G. Drin, M. Mazel, P. Clair, D. Mathieu, M. Kaczorek, et J. Temsamani, « Physico-chemical requirements for cellular uptake of pAntp peptide. Role of lipid-binding affinity », *Eur. J. Biochem. FEBS*, vol. 268, n° 5, p. 1304- 1314, mars 2001.
- [256] F. Burlina, S. Sagan, G. Bolbach, et G. Chassaing, « A direct approach to quantification of the cellular uptake of cell-penetrating peptides using MALDI-TOF mass spectrometry », *Nat. Protoc.*, vol. 1, n° 1, p. 200- 205, 2006.
- [257] C.-Y. Jiao, I. D. Alves, V. Point, S. Lavielle, S. Sagan, et G. Chassaing, « Comparing Lipid Photo-Cross-linking Efficacy of Penetratin Analogues Bearing Three Different Photoprobes: Dithienyl Ketone, Benzophenone, and Trifluoromethylaryldiazirine », *Bioconjugate Chem*, vol. 21, n° 2, p. 352- 359, 2010.
- [258] G. P. H. Dietz, « Cell-penetrating peptide technology to deliver chaperones and associated factors in diseases and basic research », *Curr. Pharm. Biotechnol.*, vol. 11, n° 2, p. 167- 174, févr. 2010.
- [259] J. S. Wadia, R. V. Stan, et S. F. Dowdy, « Transducible TAT-HA fusogenic peptide enhances escape of TAT-fusion proteins after lipid raft macropinocytosis », *Nat. Med.*, vol. 10, n° 3, p. 310- 315, mars 2004.
- [260] R. L. Juliano, R. Alam, V. Dixit, et H. M. Kang, « Cell-targeting and cell-penetrating peptides for delivery of therapeutic and imaging agents », *Wiley Interdiscip. Rev. Nanomed. Nanobiotechnol.*, vol. 1, n° 3, p. 324- 335, juin 2009.
- [261] F. Said Hassane, A. F. Saleh, R. Abes, M. J. Gait, et B. Lebleu, « Cell penetrating peptides: overview and applications to the delivery of oligonucleotides », *Cell. Mol. Life Sci. CMLS*, vol. 67, n° 5, p. 715- 726, mars 2010.
- [262] P. Boisguerin, C. Redt-Clouet, A. Franck-Miclo, S. Licheheb, J. Nargeot, S. Barrère-Lemaire, et B. Lebleu, « Systemic delivery of BH4 anti-apoptotic peptide using CPPs prevents cardiac ischemia-reperfusion injuries in vivo », *J. Control. Release Off. J. Control. Release Soc.*, vol. 156, n° 2, p. 146- 153, déc. 2011.
- [263] C.-Y. Lee, J.-F. Li, J.-S. Liou, Y.-C. Charng, Y.-W. Huang, et H.-J. Lee, « A gene delivery system for human cells mediated by both a cell-penetrating peptide and a piggyBac transposase », *Biomaterials*, vol. 32, n° 26, p. 6264- 6276, sept. 2011.
- [264] Z. Hai-Tao, L. Hui-Cheng, L. Zheng-Wu, et G. Chang-Hong, « A tumor-penetrating peptide modification enhances the antitumor activity of endostatin in vivo », *Anticancer. Drugs*, vol. 22, n° 5, p. 409- 415, juin 2011.

- [265] L. Kurzawa, M. Pellerano, J. B. Coppolani, et M. C. Morris, « Fluorescent peptide biosensor for probing the relative abundance of cyclin-dependent kinases in living cells », *PloS One*, vol. 6, n° 10, p. e26555, 2011.
- [266] V. Vukojević, A. Gräslund, et G. Bakalkin, « Fluorescence imaging with single-molecule sensitivity and fluorescence correlation spectroscopy of cell-penetrating neuropeptides », *Methods Mol. Biol. Clifton NJ*, vol. 789, p. 147- 170, 2011.
- [267] H. Nagahara, A. M. Vocero-Akbani, E. L. Snyder, A. Ho, D. G. Latham, N. A. Lissy, M. Becker-Hapak, S. A. Ezhevsky, et S. F. Dowdy, « Transduction of full-length TAT fusion proteins into mammalian cells: TAT-p27Kip1 induces cell migration », *Nat. Med.*, vol. 4, n° 12, p. 1449- 1452, déc. 1998.
- [268] S. Deshayes, K. Konate, G. Aldrian, L. Crombez, F. Heitz, et G. Divita, « Structural polymorphism of non-covalent peptide-based delivery systems: highway to cellular uptake », *Biochim. Biophys. Acta*, vol. 1798, n° 12, p. 2304- 2314, déc. 2010.
- [269] S. El-Andaloussi, P. Järver, H. J. Johansson, et U. Langel, « Cargo-dependent cytotoxicity and delivery efficacy of cell-penetrating peptides: a comparative study », *Biochem. J.*, vol. 407, n° 2, p. 285- 292, oct. 2007.
- [270] G. Tünnemann, R. M. Martin, S. Haupt, C. Patsch, F. Edenhofer, et M. C. Cardoso, « Cargo-dependent mode of uptake and bioavailability of TAT-containing proteins and peptides in living cells », *FASEB J. Off. Publ. Fed. Am. Soc. Exp. Biol.*, vol. 20, n° 11, p. 1775- 1784, sept. 2006.
- [271] O. Boussif, F. Lezoualc'h, M. A. Zanta, M. D. Mergny, D. Scherman, B. Demeneix, et J. P. Behr, « A versatile vector for gene and oligonucleotide transfer into cells in culture and in vivo: polyethylenimine », *Proc. Natl. Acad. Sci. U. S. A.*, vol. 92, n° 16, p. 7297- 7301, août 1995.
- [272] K. Splith et I. Neundorff, « Antimicrobial peptides with cell-penetrating peptide properties and vice versa », *Eur. Biophys. J. EBJ*, vol. 40, n° 4, p. 387- 397, avr. 2011.
- [273] J.-A. Spitz, V. Polard, A. Maksimenko, F. Subra, C. Baratti-Elbaz, R. Méallet-Renault, R. B. Pansu, P. Tauc, et C. Auclair, « Assessment of cellular actin dynamics by measurement of fluorescence anisotropy », *Anal. Biochem.*, vol. 367, n° 1, p. 95- 103, août 2007.
- [274] M. R. Bubb, I. Spector, B. B. Beyer, et K. M. Fosen, « Effects of jasplakinolide on the kinetics of actin polymerization. An explanation for certain in vivo observations », *J. Biol. Chem.*, vol. 275, n° 7, p. 5163- 5170, févr. 2000.
- [275] M. Kosuge, T. Takeuchi, I. Nakase, A. T. Jones, et S. Futaki, « Cellular internalization and distribution of arginine-rich peptides as a function of extracellular peptide concentration, serum, and plasma membrane associated proteoglycans », *Bioconjug. Chem.*, vol. 19, n° 3, p. 656- 664, mars 2008.
- [276] H. Kang, M. J. Bradley, B. R. McCullough, A. Pierre, E. E. Grintsevich, E. Reisler, et E. M. De La Cruz, « Identification of cation-binding sites on actin that drive polymerization and modulate bending stiffness », *Proc. Natl. Acad. Sci. U. S. A.*, vol. 109, n° 42, p. 16923- 16927, oct. 2012.
- [277] F. Boscá et M. A. Miranda, « Photosensitizing drugs containing the benzophenone chromophore », *J. Photochem. Photobiol. B*, vol. 43, n° 1, p. 1- 26, avr. 1998.
- [278] G. Dormán et G. D. Prestwich, « Benzophenone photophores in biochemistry », *Biochemistry (Mosc.)*, vol. 33, n° 19, p. 5661- 5673, mai 1994.
- [279] E. Sachon, O. Tasseau, S. Lavielle, S. Sagan, et G. Bolbach, « Isotope and affinity tags in photoreactive substance P analogues to identify the covalent linkage within the NK-1 receptor by MALDI-TOF analysis », *Anal. Chem.*, vol. 75, n° 23, p. 6536- 6543, déc. 2003.
- [280] S. Clavier, X. Du, S. Sagan, G. Bolbach, et E. Sachon, « An integrated cross-linking-MS approach to investigate cell penetrating peptides interacting partners », *EuPA Open Proteomics*, vol. 3, p. 229- 238, juin 2014.
- [281] J. Mitchell Wells et S. A. McLuckey, « Collision-Induced Dissociation (CID) of Peptides and Proteins », in *Methods in Enzymology*, vol. Volume 402, A. L. Burlingame, Éd. Academic Press, 2005, p. 148- 185.

- [282] K. Biemann, « Appendix 5. Nomenclature for peptide fragment ions (positive ions) », *Methods Enzymol.*, vol. 193, p. 886- 887, 1990.
- [283] J. E. P. Syka, J. J. Coon, M. J. Schroeder, J. Shabanowitz, et D. F. Hunt, « Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry », *Proc. Natl. Acad. Sci. U. S. A.*, vol. 101, n° 26, p. 9528- 9533, juin 2004.
- [284] R. A. Zubarev, N. L. Kelleher, et F. W. McLafferty, « Electron Capture Dissociation of Multiply Charged Protein Cations. A Nonergodic Process », *J. Am. Chem. Soc.*, vol. 120, n° 13, p. 3265- 3266, avr. 1998.
- [285] D. M. Good, M. Wirtala, G. C. McAlister, et J. J. Coon, « Performance characteristics of electron transfer dissociation mass spectrometry », *Mol. Cell. Proteomics MCP*, vol. 6, n° 11, p. 1942- 1951, nov. 2007.
- [286] D. L. Swaney, G. C. McAlister, M. Wirtala, J. C. Schwartz, J. E. P. Syka, et J. J. Coon, « A supplemental activation method for high efficiency electron transfer dissociation of doubly protonated peptide precursors », *Anal. Chem.*, vol. 79, n° 2, p. 477- 485, janv. 2007.
- [287] M. L. Lingdong Quan, « CID,ETD and HCD Fragmentation to Study Protein Post-Translational Modifications », *Mod. Chem. Appl.*, vol. 01, n° 01, 2013.
- [288] D. Mellacheruvu, Z. Wright, A. L. Couzens, J.-P. Lambert, N. A. St-Denis, T. Li, Y. V. Miteva, S. Hauri, M. E. Sardi, T. Y. Low, V. A. Halim, R. D. Bagshaw, N. C. Hubner, A. Al-Hakim, A. Bouchard, D. Faubert, D. Fermin, W. H. Dunham, M. Goudreault, Z.-Y. Lin, B. G. Badillo, T. Pawson, D. Durocher, B. Coulombe, R. Aebersold, G. Superti-Furga, J. Colinge, A. J. R. Heck, H. Choi, M. Gstaiger, S. Mohammed, I. M. Cristea, K. L. Bennett, M. P. Washburn, B. Raught, R. M. Ewing, A.-C. Gingras, et A. I. Nesvizhskii, « The CRAPome: a contaminant repository for affinity purification-mass spectrometry data », *Nat. Methods*, vol. 10, n° 8, p. 730- 736, août 2013.
- [289] H. Choi, T. Glatter, M. Gstaiger, et A. I. Nesvizhskii, « SAINT-MS1: protein-protein interaction scoring using label-free intensity data in affinity purification-mass spectrometry experiments », *J. Proteome Res.*, vol. 11, n° 4, p. 2619- 2624, avr. 2012.
- [290] A. A. Bogan et K. S. Thorn, « Anatomy of hot spots in protein interfaces », *J. Mol. Biol.*, vol. 280, n° 1, p. 1- 9, juill. 1998.
- [291] T. S. Bayer, L. N. Booth, S. M. Knudsen, et A. D. Ellington, « Arginine-rich motifs present multiple interfaces for specific binding by RNA », *RNA N. Y. N.*, vol. 11, n° 12, p. 1848- 1857, déc. 2005.
- [292] R. Petermann, B. M. Mossier, D. N. Aryee, V. Khazak, E. A. Golemis, et H. Kovar, « Oncogenic EWS-Flt1 interacts with hSRP7, a subunit of human RNA polymerase II », *Oncogene*, vol. 17, n° 5, p. 603- 610, août 1998.
- [293] J. A. Toretsky, V. Erkizan, A. Levenson, O. D. Abaan, J. D. Parvin, T. P. Cripe, A. M. Rice, S. B. Lee, et A. Uren, « Oncoprotein EWS-FLI1 activity is enhanced by RNA helicase A », *Cancer Res.*, vol. 66, n° 11, p. 5574- 5581, juin 2006.
- [294] H. V. Erkizan, Y. Kong, M. Merchant, S. Schlottmann, J. S. Barber-Rotenberg, L. Yuan, O. D. Abaan, T.-H. Chou, S. Dakshanamurthy, M. L. Brown, A. Uren, et J. A. Toretsky, « A small molecule blocking oncogenic protein EWS-FLI1 interaction with RNA helicase A inhibits growth of Ewing's sarcoma », *Nat. Med.*, vol. 15, n° 7, p. 750- 756, juill. 2009.
- [295] H. V. Erkizan, L. J. Scher, S. E. Gamble, J. S. Barber-Rotenberg, K. P. Sajwan, A. Üren, et J. A. Toretsky, « Novel peptide binds EWS-FLI1 and reduces the oncogenic potential in Ewing tumors », *Cell Cycle Georget. Tex.*, vol. 10, n° 19, p. 3397- 3408, oct. 2011.
- [296] S.-E. Ong, B. Blagoev, I. Kratchmarova, D. B. Kristensen, H. Steen, A. Pandey, et M. Mann, « Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics », *Mol. Cell. Proteomics MCP*, vol. 1, n° 5, p. 376- 386, mai 2002.
- [297] P. Pouillet, S. Carpentier, et E. Barillot, « myProMS, a web server for management and validation of mass spectrometry-based proteomic data », *Proteomics*, vol. 7, n° 15, p. 2553- 2556, août 2007.

- [298] D. Szklarczyk, A. Franceschini, M. Kuhn, M. Simonovic, A. Roth, P. Minguéz, T. Doerks, M. Stark, J. Muller, P. Bork, L. J. Jensen, et C. von Mering, « The STRING database in 2011: functional interaction networks of proteins, globally integrated and scored », *Nucleic Acids Res.*, vol. 39, n° Database issue, p. D561- 568, janv. 2011.
- [299] O. Warburg, « On the Origin of Cancer Cells », *Science*, vol. 123, n° 3191, p. 309- 314, févr. 1956.
- [300] G. L. Semenza, D. Artemov, A. Bedi, Z. Bhujwalla, K. Chiles, D. Feldser, E. Laughner, R. Ravi, J. Simons, P. Taghavi, et H. Zhong, « “The metabolism of tumours”: 70 years later », *Novartis Found. Symp.*, vol. 240, p. 251- 260; discussion 260- 264, 2001.
- [301] M. G. Vander Heiden, S. Y. Lunt, T. L. Dayton, B. P. Fiske, W. J. Israelsen, K. R. Mattaini, N. I. Vokes, G. Stephanopoulos, L. C. Cantley, C. M. Metallo, et J. W. Locasale, « Metabolic pathway alterations that support cell proliferation », *Cold Spring Harb. Symp. Quant. Biol.*, vol. 76, p. 325- 334, 2011.
- [302] J. Haynes, J. Srivastava, N. Madson, T. Wittmann, et D. L. Barber, « Dynamic actin remodeling during epithelial-mesenchymal transition depends on increased moesin expression », *Mol. Biol. Cell*, vol. 22, n° 24, p. 4750- 4764, déc. 2011.
- [303] C. Livingstone, « IGF2 and cancer », *Endocr. Relat. Cancer*, vol. 20, n° 6, p. R321- 339, déc. 2013.
- [304] L. A. Sabourin, K. Tamai, P. Seale, J. Wagner, et M. A. Rudnicki, « Caspase 3 cleavage of the Ste20-related kinase SLK releases and activates an apoptosis-inducing kinase domain and an actin-disassembling region », *Mol. Cell. Biol.*, vol. 20, n° 2, p. 684- 696, janv. 2000.
- [305] S. Wagner, T. A. Flood, P. O'Reilly, K. Hume, et L. A. Sabourin, « Association of the Ste20-like kinase (SLK) with the microtubule. Role in Rac1-mediated regulation of actin dynamics during cell adhesion and spreading », *J. Biol. Chem.*, vol. 277, n° 40, p. 37685- 37692, oct. 2002.
- [306] V. Gogvadze, B. Zhivotovsky, et S. Orrenius, « The Warburg effect and mitochondrial stability in cancer cells », *Mol. Aspects Med.*, vol. 31, n° 1, p. 60- 74, févr. 2010.
- [307] H. Pelicano, D. Carney, et P. Huang, « ROS stress in cancer cells and therapeutic implications », *Drug Resist. Updat. Rev. Comment. Antimicrob. Anticancer Chemother.*, vol. 7, n° 2, p. 97- 110, avr. 2004.
- [308] A. M. Janssen, C. B. Bosman, L. Kruidenier, G. Griffioen, C. B. Lamers, J. H. van Krieken, C. J. van de Velde, et H. W. Verspaget, « Superoxide dismutases in the human colorectal cancer sequence », *J. Cancer Res. Clin. Oncol.*, vol. 125, n° 6, p. 327- 335, 1999.
- [309] K. Punnonen, M. Ahotupa, K. Asaishi, M. Hyöty, R. Kudo, et R. Punnonen, « Antioxidant enzyme activities and oxidative stress in human breast cancer », *J. Cancer Res. Clin. Oncol.*, vol. 120, n° 6, p. 374- 377, 1994.
- [310] L. W. Oberley, I. B. Bize, S. K. Sahu, S. W. Leuthauser, et H. E. Gruber, « Superoxide dismutase activity of normal murine liver, regenerating liver, and H6 hepatoma », *J. Natl. Cancer Inst.*, vol. 61, n° 2, p. 375- 379, août 1978.
- [311] B. E. Van Driel, H. Lyon, D. C. Hoogenraad, S. Anten, U. Hansen, et C. J. Van Noorden, « Expression of CuZn- and Mn-superoxide dismutase in human colorectal neoplasms », *Free Radic. Biol. Med.*, vol. 23, n° 3, p. 435- 444, 1997.
- [312] T. G. Graeber, C. Osmanian, T. Jacks, D. E. Housman, C. J. Koch, S. W. Lowe, et A. J. Giaccia, « Hypoxia-mediated selection of cells with diminished apoptotic potential in solid tumours », *Nature*, vol. 379, n° 6560, p. 88- 91, janv. 1996.
- [313] R. A. Gatenby et R. J. Gillies, « Why do cancers have high aerobic glycolysis? », *Nat. Rev. Cancer*, vol. 4, n° 11, p. 891- 899, nov. 2004.
- [314] H. Zhang, X. Chen, W. B. Bollag, R. J. Bollag, D. J. Sheehan, et C. S. Chew, « Lasp1 gene disruption is linked to enhanced cell migration and tumor formation », *Physiol. Genomics*, vol. 38, n° 3, p. 372- 385, août 2009.
- [315] Y. Estornes, F. Gay, J.-C. Gevrey, S. Navoizat, M. Nejari, J.-Y. Scoazec, J.-A. Chayvialle, J.-C. Saurin, et J. Abello, « Differential involvement of destrin and cofilin-1 in the control of invasive

- properties of Isreco1 human colon cancer cells », *Int. J. Cancer J. Int. Cancer*, vol. 121, n° 10, p. 2162- 2171, nov. 2007.
- [316] H. Haase, I. Pagel, Y. Khalina, U. Zacharzowsky, V. Person, G. Lutsch, D. Petzhold, M. Kott, J. Schaper, et I. Morano, « The carboxyl-terminal ahnak domain induces actin bundling and stabilizes muscle contraction », *FASEB J. Off. Publ. Fed. Am. Soc. Exp. Biol.*, vol. 18, n° 7, p. 839- 841, mai 2004.
- [317] I. Park, C. Han, S. Jin, B. Lee, H. Choi, J. T. Kwon, D. Kim, J. Kim, E. Lifirsu, W. J. Park, Z. Y. Park, D. H. Kim, et C. Cho, « Myosin regulatory light chains are required to maintain the stability of myosin II and cellular integrity », *Biochem. J.*, vol. 434, n° 1, p. 171- 180, févr. 2011.
- [318] E. Daimon, Y. Shibukawa, et Y. Wada, « Calponin 3 regulates stress fiber formation in dermal fibroblasts during wound healing », *Arch. Dermatol. Res.*, vol. 305, n° 7, p. 571- 584, sept. 2013.
- [319] M. Wilhelm, J. Schlegl, H. Hahne, A. Moghaddas Gholami, M. Lieberenz, M. M. Savitski, E. Ziegler, L. Butzmann, S. Gessulat, H. Marx, T. Mathieson, S. Lemeer, K. Schnatbaum, U. Reimer, H. Wenschuh, M. Mollenhauer, J. Slotta-Huspenina, J.-H. Boese, M. Bantscheff, A. Gerstmair, F. Faerber, et B. Kuster, « Mass-spectrometry-based draft of the human proteome », *Nature*, vol. 509, n° 7502, p. 582- 587, mai 2014.
- [320] J. Chenau, S. Michelland, J. Sidibe, et M. Seve, « Peptides OFFGEL electrophoresis: a suitable pre-analytical step for complex eukaryotic samples fractionation compatible with quantitative iTRAQ labeling », *Proteome Sci.*, vol. 6, p. 9, 2008.
- [321] K. C. Chan et H. J. Issaq, « Fractionation of peptides by strong cation-exchange liquid chromatography », *Methods Mol. Biol. Clifton NJ*, vol. 1002, p. 311- 315, 2013.
- [322] R. A. Cairns, I. S. Harris, et T. W. Mak, « Regulation of cancer cell metabolism », *Nat. Rev. Cancer*, vol. 11, n° 2, p. 85- 95, févr. 2011.
- [323] X. Du, S. M. Chowdhury, N. P. Manes, S. Wu, M. U. Mayer, J. N. Adkins, G. A. Anderson, et R. D. Smith, « Xlink-identifier: an automated data analysis platform for confident identifications of chemically cross-linked peptides using tandem mass spectrometry », *J. Proteome Res.*, vol. 10, n° 3, p. 923- 931, mars 2011.
- [324] J. R. Wiśniewski, A. Zougman, N. Nagaraj, et M. Mann, « Universal sample preparation method for proteome analysis », *Nat. Methods*, vol. 6, n° 5, p. 359- 362, mai 2009.
- [325] M. Karas et F. Hillenkamp, « Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons », *Anal. Chem.*, vol. 60, n° 20, p. 2299- 2301, oct. 1988.
- [326] K. Tanaka, H. Waki, Y. Ido, S. Akita, Y. Yoshida, T. Yoshida, et T. Matsuo, « Protein and polymer analyses up to m/z 100 000 by laser ionization time-of-flight mass spectrometry », *Rapid Commun. Mass Spectrom.*, vol. 2, n° 8, p. 151- 153, août 1988.
- [327] J. B. Fenn, M. Mann, C. K. Meng, S. F. Wong, et C. M. Whitehouse, « Electrospray ionization for mass spectrometry of large biomolecules », *Science*, vol. 246, n° 4926, p. 64- 71, oct. 1989.
- [328] N. B. Cech et C. G. Enke, « Practical implications of some recent studies in electrospray ionization fundamentals », *Mass Spectrom. Rev.*, vol. 20, n° 6, p. 362- 387, déc. 2001.
- [329] null Makarov, « Electrostatic axially harmonic orbital trapping: a high-performance technique of mass analysis », *Anal. Chem.*, vol. 72, n° 6, p. 1156- 1162, mars 2000.
- [330] Q. Hu, R. J. Noll, H. Li, A. Makarov, M. Hardman, et R. Graham Cooks, « The Orbitrap: a new mass spectrometer », *J. Mass Spectrom. JMS*, vol. 40, n° 4, p. 430- 443, avr. 2005.



# Annexes





## I. Annexe instrumentale : Spectrométrie de masse

Les instruments dont lesquels j'ai travaillé sont le spectromètre de masse MALDI-TOF/TOF 4700 Proteomics Analyzer (AB Sciex) et un système couplant la chromatographie liquide à très bas débit avec la spectrométrie de masse, le système NanoLC (Ultimate 3000, Dionex)-ESI-MS/MS (LTQ-Orbitrap XL, Thermo Scientific). Ces deux instruments se sont avérés très complémentaires dans la réalisation de ces travaux de thèse. Le système MALDI-TOF/TOF était essentiellement utilisé en première intention de manière à obtenir rapidement une idée du contenu des échantillons tandis que le couplage NanoLC-ESI-LTQ-Orbitrap s'est révélé indispensable pour réaliser l'analyse haut débit d'échantillons biologiques complexes (lysats cellulaires et milieux issus d'expérience de purification par affinité). De plus, comme détaillé dans le chapitre 5, nous avons pu comparer les modes de fragmentation disponibles sur ces instruments, CID « haute énergie » sur le MALDI-TOF/TOF, CID « basse énergie » et ETD sur le LTQ-Orbitrap XL, pour l'analyse des peptides photocross-linkés.

Pour fixer un peu mieux les choses, en particulier pour les non-massistes, se trouve ci-après une brève description des techniques d'ionisation et des analyseurs utilisés. Finalement, les schémas des instruments seront donnés et le devenir des ions dans ces spectromètres de masse sera détaillé.

### 1. Spectromètres de masse : Généralités

Tous les spectromètres de masse sont structurés de la manière suivante : une source d'ions, un analyseur et un détecteur comme indiqué sur le schéma ci-dessous (Figure annexe 1). Le signal obtenu est ensuite traité pour donner un spectre.

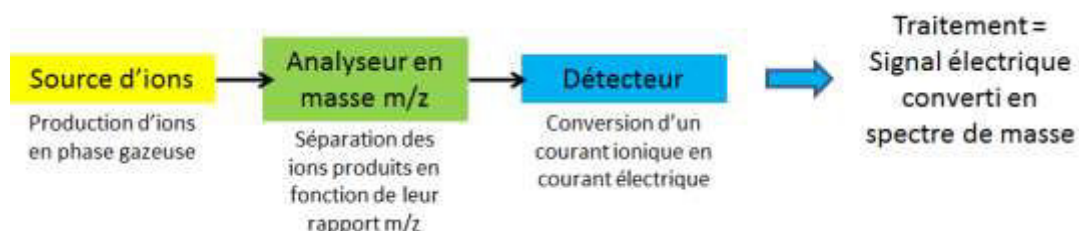


Figure annexe 1: Structure d'un spectromètre de masse.

Les méthodes d'ionisation utilisées dans ce travail sont **l'électronébulisation** notée **ESI** et **la désorption/ionisation laser assistée par matrice** notée **MALDI**. Ce sont des méthodes d'ionisation qualifiées de « douces » qui permettent de transformer des molécules initialement neutres électriquement dans un état liquide (ESI) ou solide (MALDI) en molécules chargées électriquement (positif ou négatif) et en phase gazeuse en induisant peu ou pas de fragmentation de ces molécules. Ce sont ces méthodes d'ionisation qui sont ainsi principalement utilisées pour l'analyse de biomolécules.

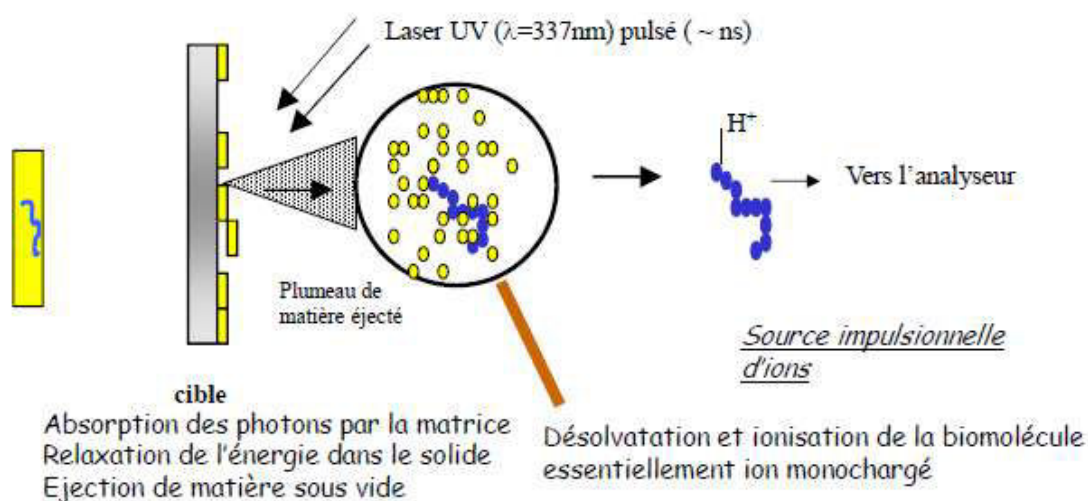
Concernant les analyseurs utilisés, ils sont au nombre de trois : l'analyseur temps de vol ou TOF (*time of flight*), le piège ionique linéaire ou LTQ (*Linear trap quadrupole*) et l'Orbitrap.

Dans le cas des spectromètres de masse MALDI-TOF/TOF et LTQ, la détection est réalisée par des multiplicateurs d'électrons après conversion ion-électron (galettes de microcanaux en TOF, dynode continue en LTQ) tandis qu'elle se fait directement dans « l'analyseur » dans le cas de l'Orbitrap par mesure de courant image.

#### a. Sources d'ions ESI et MALDI :

##### **Désorption/ionisation laser assistée par matrice (MALDI) ([325], [326])**

La désorption/ionisation laser assistée par matrice (MALDI) est, comme son nom l'indique, un mode de production d'ions au moyen d'un faisceau laser. Les biomolécules à analyser sont incorporées dans des cristaux de matrice formée de petites molécules organiques absorbant à la longueur d'onde du laser (ex :  $\lambda = 337 \text{ nm}$ ). Cette **absorption des photons UV** par les molécules de matrice va être à l'origine d'onde de pression dans le matériau avec l'éjection et l'ionisation de la matrice ( $m^{++}$ ,  $mH^+$ ). Cette éjection se traduit par l'expulsion **d'un plumeau de matière**. Dans ce plumeau de matière éjectée, des agrégats composés de molécules de matrice (neutres et ionisées sous forme protonées) et de biomolécules, des réactions ions molécules vont conduire à des transferts de protons de la matrice vers les biomolécules (cf. Figure annexe 2). La source d'ions MALDI est une **source impulsienne**, les ions ne sont donc pas formés de manière continue mais suite à chaque impulsion laser.



**Figure annexe 2 : Schéma de principe de la production d'ions par MALDI**

En MALDI, le nombre de protons transférés reste limité. Ainsi les peptides (MW < 5000 Da) vont être majoritairement monochargés tandis qu'une protéine de 60 kDa sera présente avec des états de charges typiquement compris entre 1 et 10 maximum. L'ion monochargé est toujours observé quel que soit la taille de l'analyte.

##### **Electronébulisation (Electrospray Ionisation, ESI) :**

L'électronébulisation introduite pour la première fois en 1988 ([327]) repose sur l'introduction d'une solution aqueuse du composé à analyser dans un capillaire très fin porté à un haut potentiel (~kV). A l'extrémité libre du capillaire cette tension crée une séparation de charges dans la solution en

formant un **nébulisat (« spray ») de fines gouttelettes chargées ou non** (cf. Figure annexe 3). Dans la source à pression atmosphérique, ces gouttelettes chauffées vont voir leur taille diminuer progressivement par évaporation du solvant (désolvatation). La diminution de la taille des gouttelettes chargées va conduire également à une augmentation de la densité de charges au sein de la goutte. Lorsque la répulsion coulombienne excède la tension superficielle de la gouttelette (limite de Rayleigh), celle-ci explose en formant des gouttelettes de seconde génération, beaucoup plus petites. Finalement cela conduit à la production d'ions en phase gazeuse, physiquement intacts et pouvant porter **plusieurs charges** en raison de l'attachement de protons aux sites basiques des molécules. La source d'ions ESI est une **source continue** tant que le débit de liquide et la tension sont maintenus. Les ions obtenus, majoritairement multichargés lorsqu'il s'agit de peptides, vont être transférés dans l'analyseur grâce à la différence de potentiel appliquée entre le capillaire métallisé et l'entrée de l'analyseur. Typiquement pour une protéine dénaturée, une charge peut être fixée pour une masse de ~1000 Da soit pour une chaîne peptidique de 9 à 10 acides aminés.

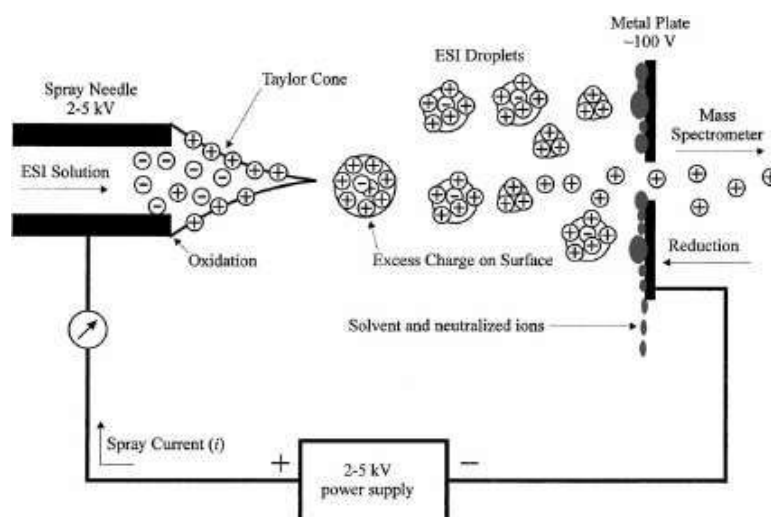


Figure annexe 3 : Schéma du processus d'ionisation par électronébulisation (figure tirée de [328]).

## b. Analyseurs utilisés

### L'analyseur temps de vol (TOF, Time of flight)

L'analyseur **temps de vol** (TOF) est dans son principe le plus simple de tous les analyseurs. Il est parfaitement adapté aux sources d'ions impulsionnelles pour lesquelles tous les ions sont produits pendant un temps très court (de l'ordre de la nanoseconde) comme en MALDI.

Les ions créés dans le voisinage de la cible MALDI suite à l'éjection de matière sont accélérés par un champ électrique homogène de quelques kV appliqué entre la cible et une grille (cf. Figure annexe 4). L'énergie cinétique acquise par les ions de masse  $m$  et de charge  $ze$  possédant une énergie potentielle  $zeV$  sera égale à  $0.5 Mv^2 = zeV$  en atteignant la grille soit  $v = (2zeV/M)^{1/2}$ . Ainsi plus la masse de l'ion sera élevée plus la vitesse sera faible et pour une molécule de masse  $m$ , les ions doublement chargés seront plus accélérés que les ions monochargés. Ayant traversé la grille, les ions vont entrer dans une région sans champ électrique et vont poursuivre leur chemin avec la vitesse acquise au cours de la phase d'accélération jusqu'à atteindre le détecteur. La **différence de vitesse**

des ions entraînera un **décalage dans le temps d'arrivée au détecteur**. Les ions sont donc distingués par leur différence de temps d'arrivée au détecteur. Une propriété intéressante du TOF est que l'ensemble des ions produits par un tir MALDI sera détecté et donc analysé.

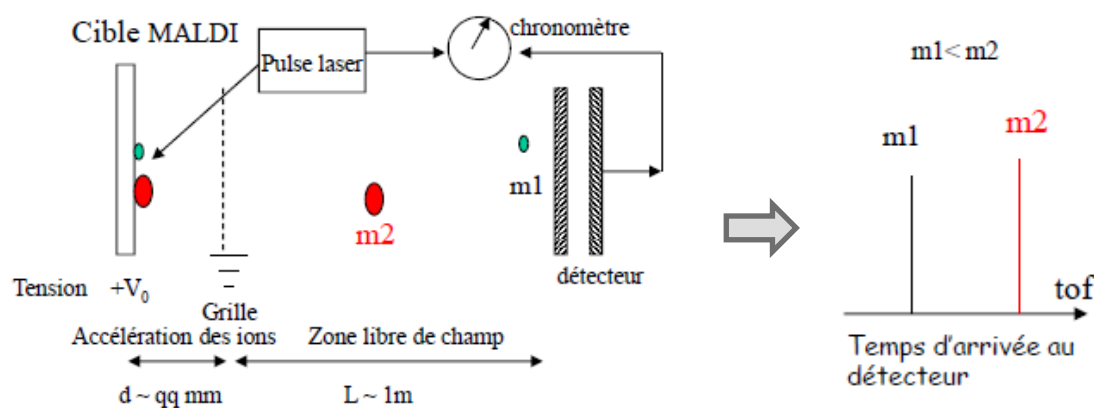


Figure annexe 4 : Principe de la séparation des ions et de la mesure de leur rapport  $m/z$  avec un analyseur à temps de vol (TOF).

L'analyseur TOF présente l'avantage d'avoir une gamme de  $m/z$  théoriquement illimitée (Une fois qu'une biomolécule de taille importante a pu être correctement ionisée, il « suffit d'attendre suffisamment longtemps » pour la détecter) et est donc très adapté à l'analyse de protéines entières.

Outre l'analyse en **mode linéaire** pour laquelle la détection intervient à l'issue d'un aller simple dans le tube de vol, il est possible d'utiliser le mode réflecteur. Le **mode « reflectron »** repose sur l'utilisation d'un miroir électrostatique qui va venir corriger les écarts de vitesse d'ions de même  $m/z$  et peut être appliqué à des ions avec des masses jusqu'à 4000-5000 Da (cf. Figure annexe 5). Au-delà, les ions sont trop lents pour être correctement réfléchis. En passant du mode linéaire au mode réflecteur on augmente considérablement la résolution (de quelques milliers jusqu'à environ 20.000 de résolution).

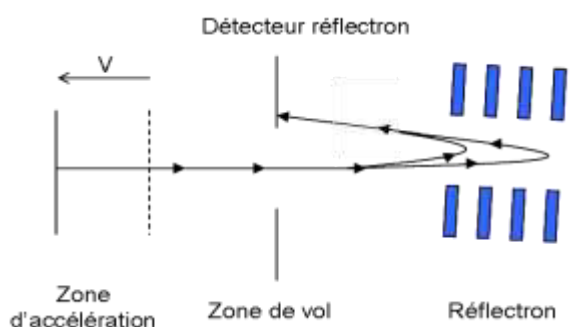
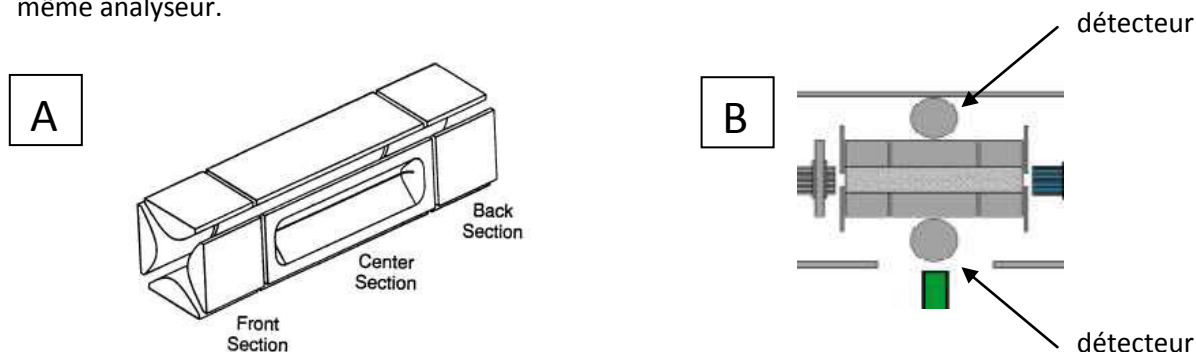


Figure annexe 5 : Schéma de la réflexion électrostatique des ions dans un analyseur TOF en mode réflecteur

#### L'analyseur LTQ, trappe ionique linéaire

L'analyseur LTQ fait partie des **analyseurs à stabilité de trajectoire (quadripôle, piège quadripolaire)**. Dans ces analyseurs règne un champ électrique oscillant de fréquence  $\nu_{RF}$  ( $\sim \text{MHz}$ ) qui va donner aux ions qui y sont introduits une trajectoire radiale stable ou instable suivant leur rapport  $m/z$ . Un ion de trajectoire instable frappera la surface des électrodes et sera perdu. Un ion de

trajectoire stable sera piégé et sélectionné. Dans le piège ionique linéaire, les ions sont piégés radialement par le champ radio-fréquence RF et piégés axialement par des tensions appliquées aux extrémités du quadripôle (cf. Figure annexe 6.A). Dans ce piège linéaire, il est possible d'accumuler un plus grand nombre d'ions injectés de la source ESI et donc d'augmenter la sensibilité analytique. L'augmentation de la tension RF déstabilise radialement les ions les dirigeant vers les deux détecteurs (cf. Figure annexe 6.B) La sélection d'un ion et sa fragmentation CID se produisent dans le même analyseur.



Le piège linéaire quadripolaire (LTQ) permet de travailler sur une gamme de masse allant de 15 à 4000  $m/z$  mais elle est classiquement utilisée sur une gamme plus restreinte (ex : 500-2000  $m/z$ ), de manière à diminuer la durée de chaque cycle. La résolution d'une trappe ionique est de l'ordre de 5000.

### Analyseur Orbitrap

L'analyseur **Orbitrap** inventé par A. Makarov en 2000 ([329]) et commercialisé pour la première fois par Thermo Scientific en 2005 ([330]) est basé sur un **piégeage orbital des ions** par application d'un potentiel électrique unique constant sur une électrode centrale en forme de fuseau (cf. Figure annexe 7). Les ions injectés (via une optique adaptée C-trap) sont piégés radialement et axialement. L'électrode extérieure scindée en deux parties isolées électriquement permet de détecter le passage des ions par un effet de charge-image. La mesure de la fréquence du mouvement périodique axial (suivant  $z$ ) permet de remonter au rapport  $m/z$ . Cet analyseur permet d'atteindre une résolution importante (supérieure à 100.000) et une très bonne précision (quelques ppm) avec beaucoup moins de contraintes que les spectromètres de masse FTICR (coût, maintenance, fonctionnement...).

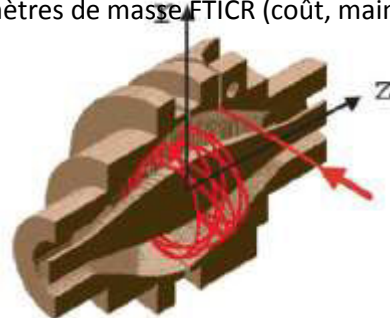


Figure annexe 7 : Vue en coupe de l'analyseur Orbitrap. Les ions sont introduits dans l'analyseur au niveau de la flèche rouge (Figure extraite de[330]).

## 2. Représentations schématiques des spectromètres de masse utilisés

Après le rappel de ces généralités concernant les sources d'ions et les analyseurs, intéressons-nous aux représentations schématiques des spectromètres de masse utilisés, à savoir le MALDI-TOF/TOF 4700 Proteomics Analyzer (cf. Figure annexe 8) et le spectromètre de masse LTQ-Orbitrap XL (cf. Figure annexe 9).

### a. Spectromètre de masse MALDI-TOF/TOF (4700 Proteomics Analyzer)

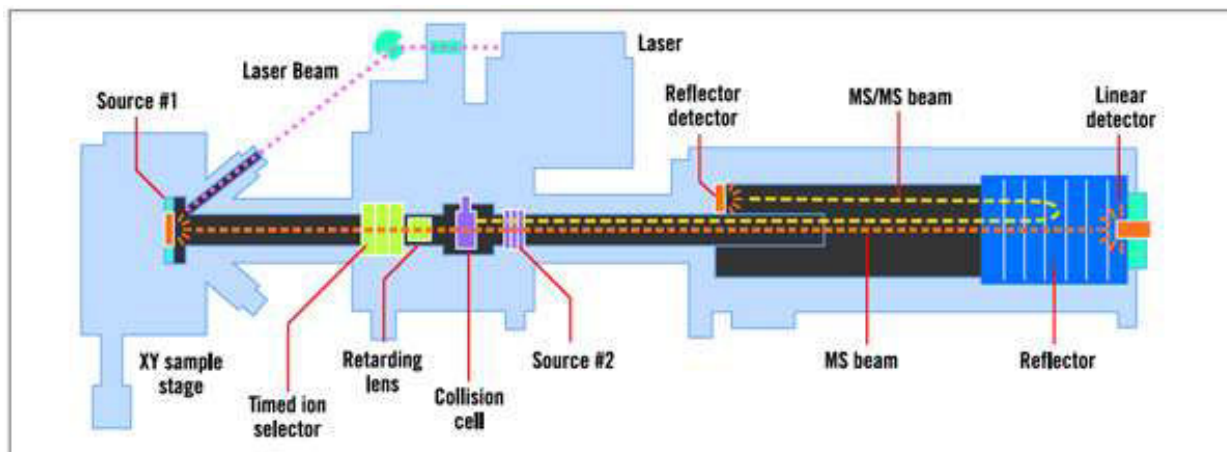


Figure annexe 8 : Représentation schématique du spectromètre de masse MALDI-TOF/TOF 4700 Proteomics Analyzer

Le spectromètre de masse MALDI-TOF/TOF 4700 Proteomics Analyzer est composé d'une **chambre d'ionisation** placée sous vide dans laquelle va être introduite la plaque MALDI sur laquelle les échantillons auront été cocrystallisés avec de la matrice (classiquement CHCA). Les différentes cibles seront successivement analysées en déplaçant la plaque MALDI par rapport au faisceau laser impulsif ( $\lambda = 355 \text{ nm}$ , fréquence de répétition 200 Hz). Après la chambre d'ionisation se trouve un dispositif noté « **retarding lens** » qui permet, par un processus dit **d'extraction retardée**, de compenser les écarts de vitesse initiale des ions générés. Ce spectromètre de masse est également équipé d'un **miroir électrostatique** (en bleu à droite de l'image) destiné comme indiqué précédemment (I.2.a) à augmenter la résolution en compensant les écarts de vitesse initiale. Les ions peuvent être analysés sous forme intacte directement en mode linéaire ou en mode réflecteur.

Il est également possible de **fragmenter les ions produits** en source dans une cellule de collision au moyen d'un gaz (l'air dans nos études) afin d'obtenir des informations structurales sur les analytes. Après une sélection d'un ion particulier (TOF), les fragments obtenus par collision énergétique (1keV) sont analysés par un second TOF et détectés en mode réflecteur.



## b. Spectromètre de masse LTQ-Orbitrap XL

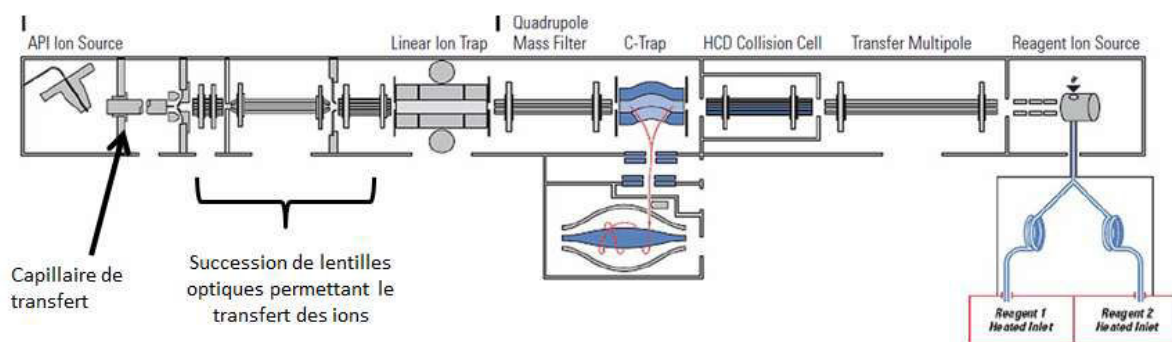


Figure annexe 9 : Schéma du spectromètre de masse LTQ-Orbitrap XL (Thermo Scientific)

Sur ce spectromètre de masse, il est possible d'adapter différentes sources : ESI, nanoESI et MALDI. C'est avec la **source nanoESI** et des débits compris entre 200 et 300nl/min que nous avons travaillé. Une fois les analytes introduits en solution via la source nanoESI, ils sont ionisés après désolvatation des gouttelettes en traversant un capillaire de transfert chauffé. Ils entrent ensuite dans le spectromètre de masse. Grâce à une série de lentilles optiques, ils sont ensuite transférés dans la piège ionique linéaire (LTQ). Ces ions peuvent être directement analysés dans le LTQ (MS ou MS/MS) ou bien envoyés vers l'Orbitrap pour une détermination plus précise et avec une meilleure résolution des rapports  $m/z$  des ions. Au-delà de l'analyse MS des ions, ce spectromètre de masse permet de fragmenter les ions de trois manières : en CID directement dans le LTQ, en HCD (CID avec un peu plus d'énergie) dans la cellule de collision HCD) et en ETD. (NB : Nous avons essentiellement travaillé avec les modes CID et ETD car la transmission des fragments en HCD n'était pas satisfaisante). Pour davantage de détails concernant les modes de fragmentation CID et ETD se référer au chapitre 5. Partie II.2.c. Les fragments obtenus peuvent être analysés soit dans le LTQ soit en haute résolution dans l'Orbitrap. De manière à focaliser les ions parents ou fils avant de les transférer dans l'Orbitrap, le système est muni d'une optique adaptée (C-trap).

Classiquement, nous avons réalisé la mesure des  $m/z$  des ions parents dans l'Orbitrap et celle des fragments dans le LTQ.



## II. Annexes diverses

### Chapitre 5 : Photocross-linking

**Annexe II 1:** Recherches des peptides photocross-linkés (R/W)<sub>9</sub>-BSA avec Mascot sur P.D 1.3

Recherche avec acide glutamique (E)

Workflow Editor (BSA photocross-linking search on tryptic)									
Proteins		Search Input		Result Files		Peptide Confidence		Search Summary	
#	Protein	Seq	Sequence	# Peptides	# Proteins	# Peptides	# Proteins	# Peptides	# Proteins
1	BSA	1	HPHVAISYLIR	1	2	1	2	1	2
2	BSA	2	ETKATLLEQAK	1	1	1	1	1	1
3	BSA	3	DOPHAKSTYTK	1	1	1	1	1	1
4	BSA	4	ECHGDLLEQAK	4	10	1	10	1	10
5	BSA	5	HPHVAISYLIR	3	2	1	2	1	2
6	BSA	6	ECHGDLLEQAK	1	2	1	2	1	2
7	BSA	7	ECHGDLLEQAK	2	10	1	10	1	10
8	BSA	8	ECHGDLLEQAK	3	10	1	10	1	10

Recherche avec histidine (H)

Workflow Editor (BSA photocross-linking search on tryptic)									
Proteins		Search Input		Result Files		Peptide Confidence		Search Summary	
#	Protein	Seq	Sequence	# Peptides	# Proteins	# Peptides	# Proteins	# Peptides	# Proteins
1	BSA	1	HPHVAISYLIR	1	2	1	2	1	2
2	BSA	2	ETKATLLEQAK	1	1	1	1	1	1
3	BSA	3	DOPHAKSTYTK	1	1	1	1	1	1
4	BSA	4	ECHGDLLEQAK	4	10	1	10	1	10
5	BSA	5	HPHVAISYLIR	3	2	1	2	1	2
6	BSA	6	ECHGDLLEQAK	1	2	1	2	1	2
7	BSA	7	ECHGDLLEQAK	2	10	1	10	1	10
8	BSA	8	ECHGDLLEQAK	3	10	1	10	1	10

Recherche avec thréonine (T)

Workflow Editor (BSA photocross-linking search on tryptic)									
Proteins		Search Input		Result Files		Peptide Confidence		Search Summary	
#	Protein	Seq	Sequence	# Peptides	# Proteins	# Peptides	# Proteins	# Peptides	# Proteins
1	BSA	1	HPHVAISYLIR	1	2	1	2	1	2
2	BSA	2	ETKATLLEQAK	1	1	1	1	1	1
3	BSA	3	DOPHAKSTYTK	1	1	1	1	1	1
4	BSA	4	ECHGDLLEQAK	4	10	1	10	1	10
5	BSA	5	HPHVAISYLIR	3	2	1	2	1	2
6	BSA	6	ECHGDLLEQAK	1	2	1	2	1	2
7	BSA	7	ECHGDLLEQAK	2	10	1	10	1	10
8	BSA	8	ECHGDLLEQAK	3	10	1	10	1	10

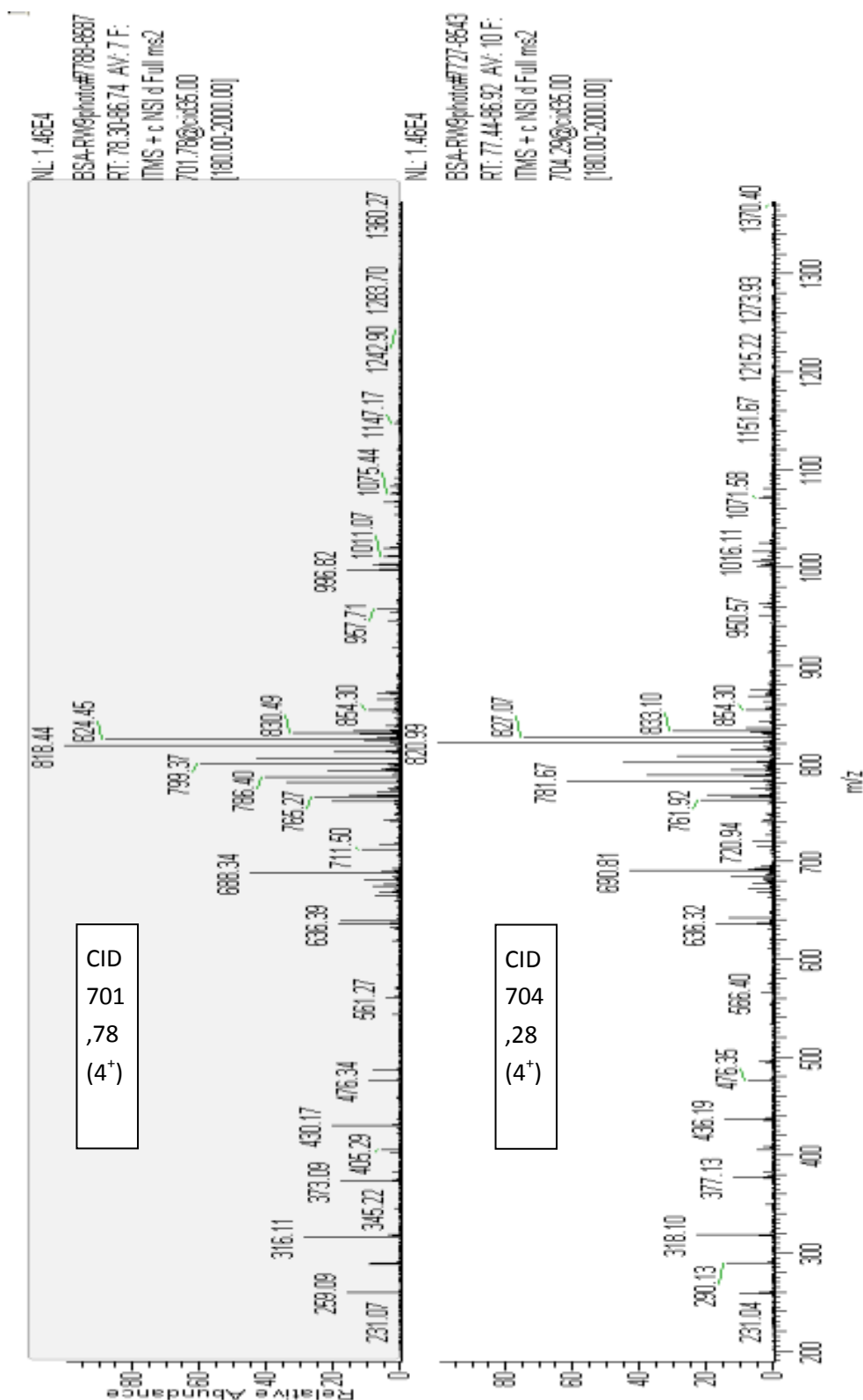
## Recherche avec tyrosine (Y)

Protein	Peptide	Search Input	Result Filters	Peptide Confidence	Search Summary			Protein Group Accessions	Modifications	ΔCn	IonScore	Exp Value	Charge	MH+ [Da]	ΔM [ppm]	RT [min]
					# PSMs	# Proteins	# Protein Groups									
+	1	hPEAVSVLLR	Sequence	1	2	1	P02769			0.0000	20	1.4E-001	2	12837.1306	1.85	33.94
+	2	EYEAITLECAK		1	1	1	P02769	C9(Carbamidomethyl); C10(Carbamidomethyl)		0.0000	57	8.0E-006	2	1502.61528	0.87	29.20
+	3	GGPHACISTVDFK		1	1	1	P02769	C6(Carbamidomethyl)		0.0000	30	5.9E-003	2	1354.45471	1.11	31.28
+	4	ECHGDLLEADRS		4	10	1	P02769	C2(Carbamidomethyl); C3(Carbamidomethyl); C10(Carbamidomethyl)		0.0000	29	1.6E-003	3	1749.56446	1.06	28.94
+	5	HPVFAPPELLYANK		3	2	1	P02769			0.0000	44	1.7E-003	3	1888.92624	-0.33	37.72
+	6	ETVGGHADCCAK		1	2	1	P02769	Y3(R49 p); C6(Carbamidomethyl); C10(Carbamidomethyl)		0.0000	35	2.0E-002	3	2531.96262	2.17	31.57

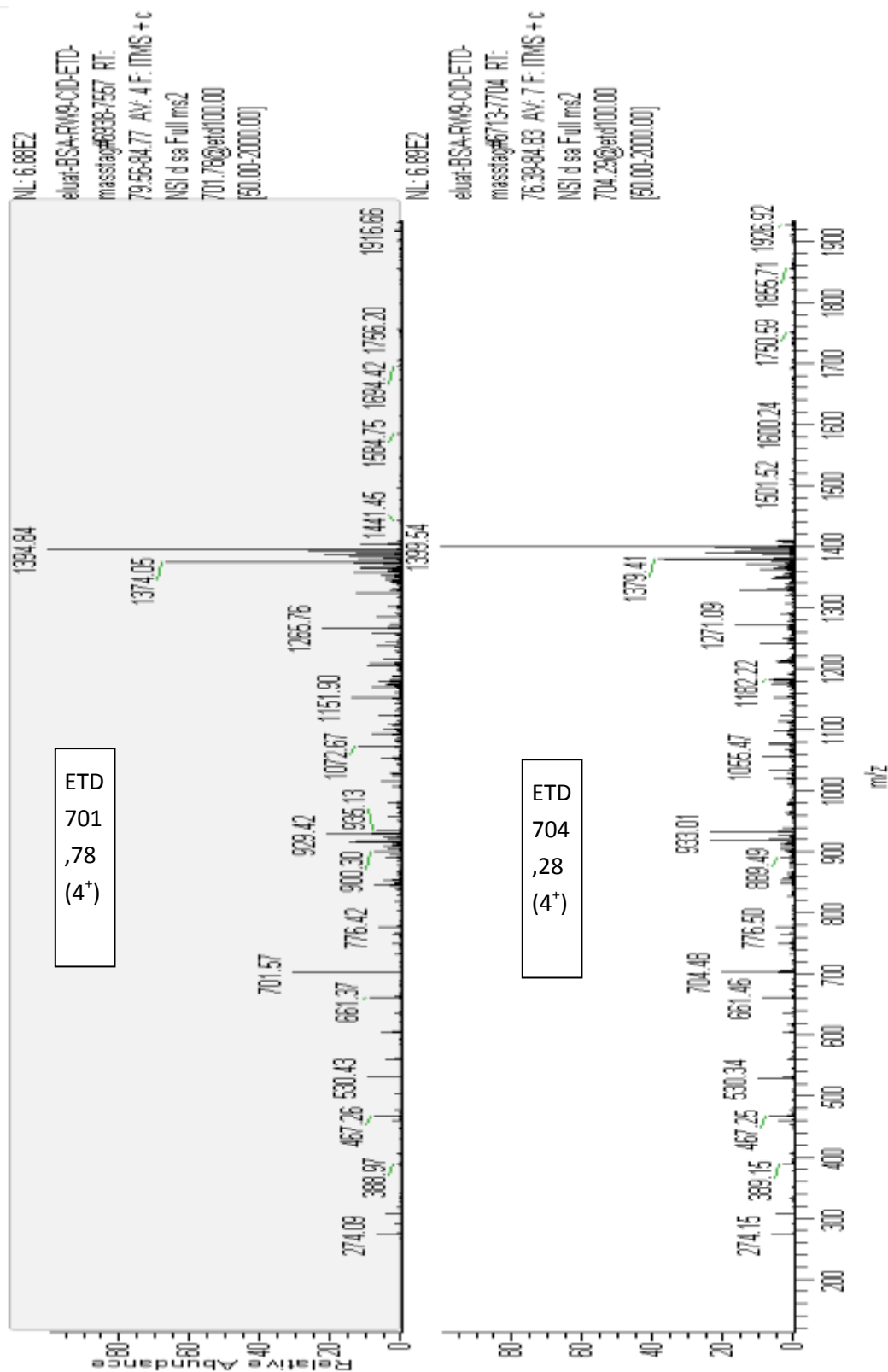
## Recherche avec leucine (L)

[illegible]

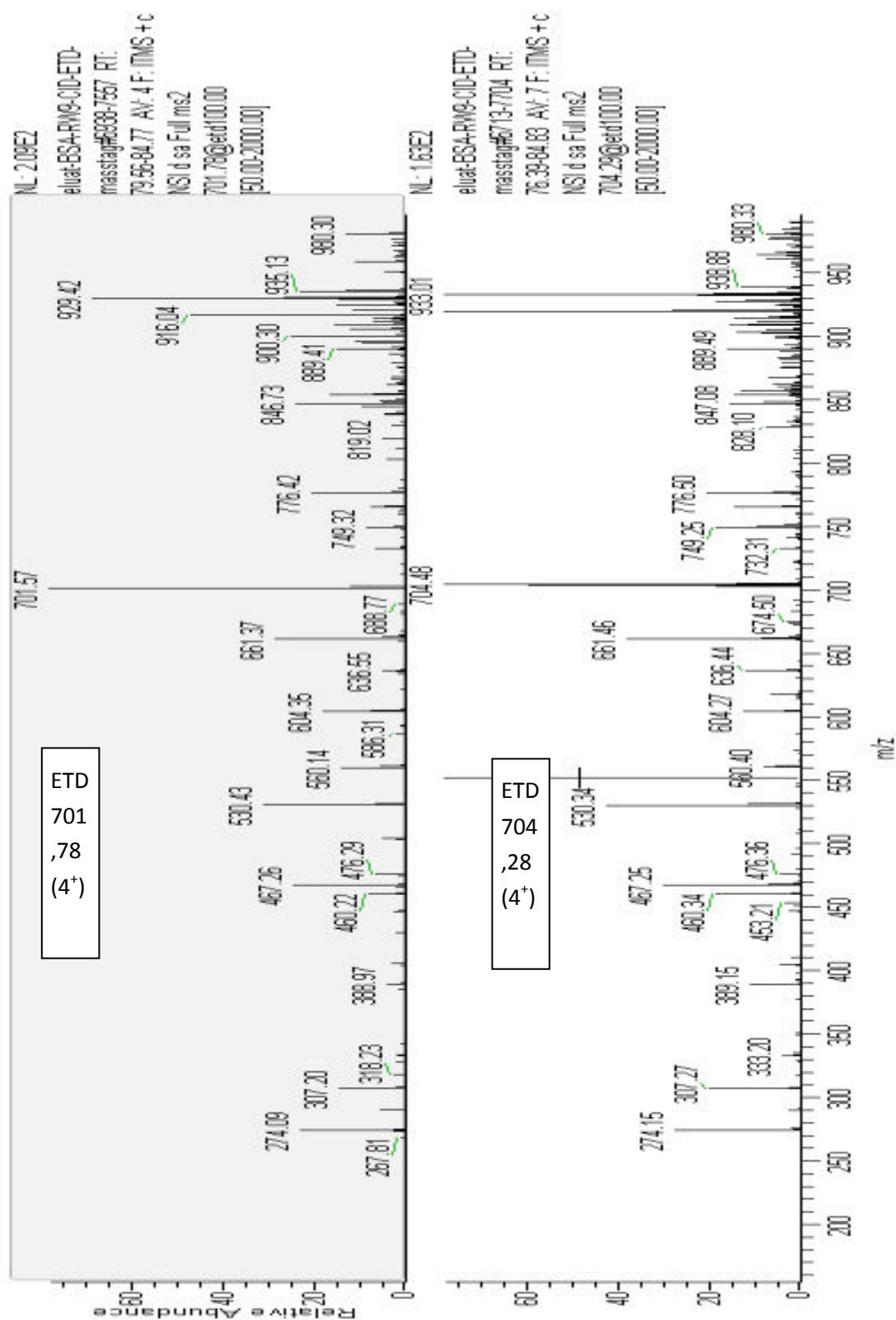
**Annexe II 2:** Spectres de fragmentation CID (basse et haute énergie) et ETD des ions de  $(M+H)^+$  2803-2813



Spectres CID basse énergie (ESI-LTQ/Orbitrap) des précurseurs 4<sup>+</sup> (701,78 et 704,29) pour les ions de  $(M+H)^+$  2803 et 2813.

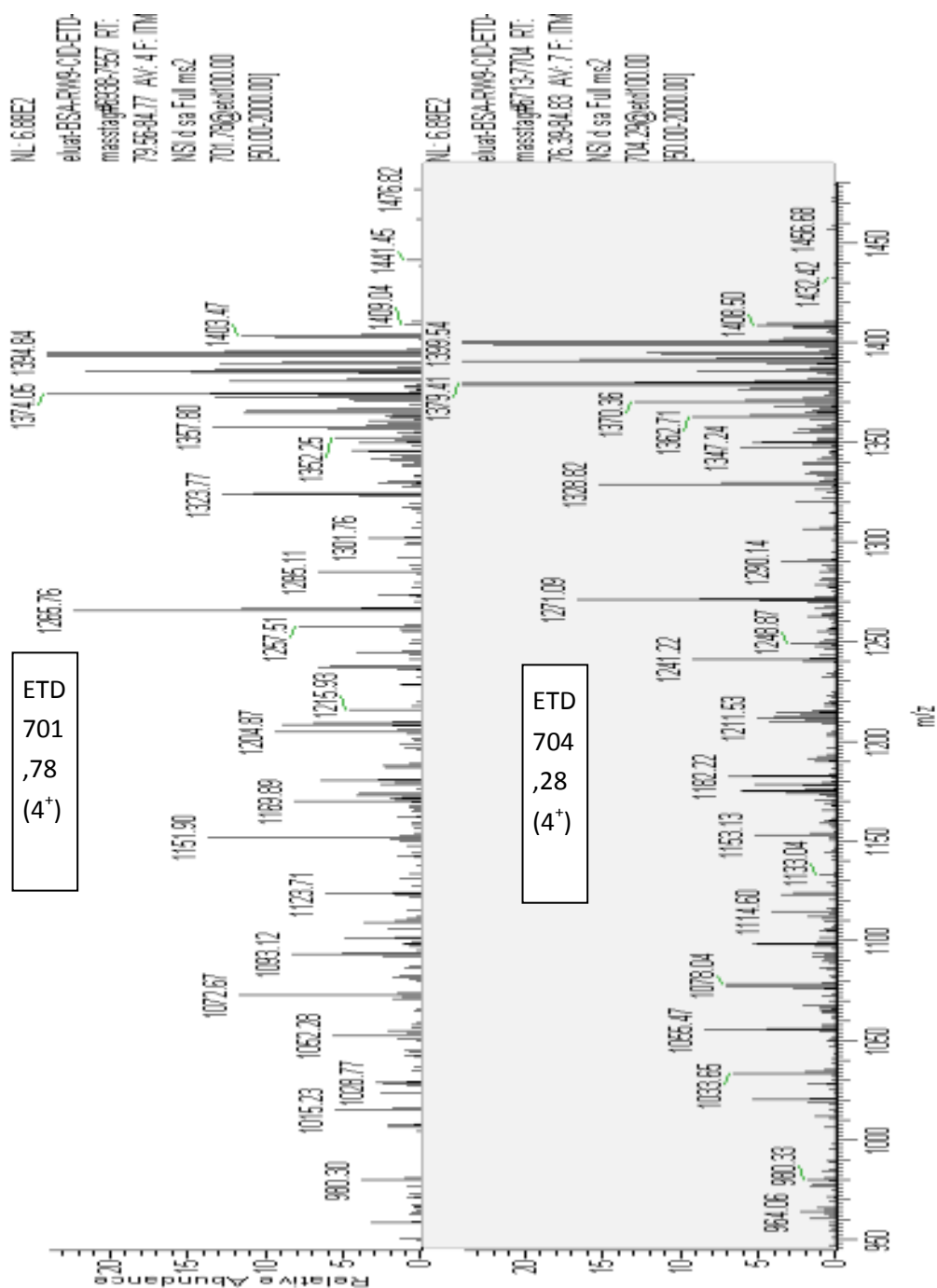


Spectres ETD (Gamme de m/z 50-2000) (ESI-LTQ/Orbitrap) des précurseurs 4+ (701,78 et 704,29) pour les ions de (M+H)<sup>+</sup> 2803 et 2813.

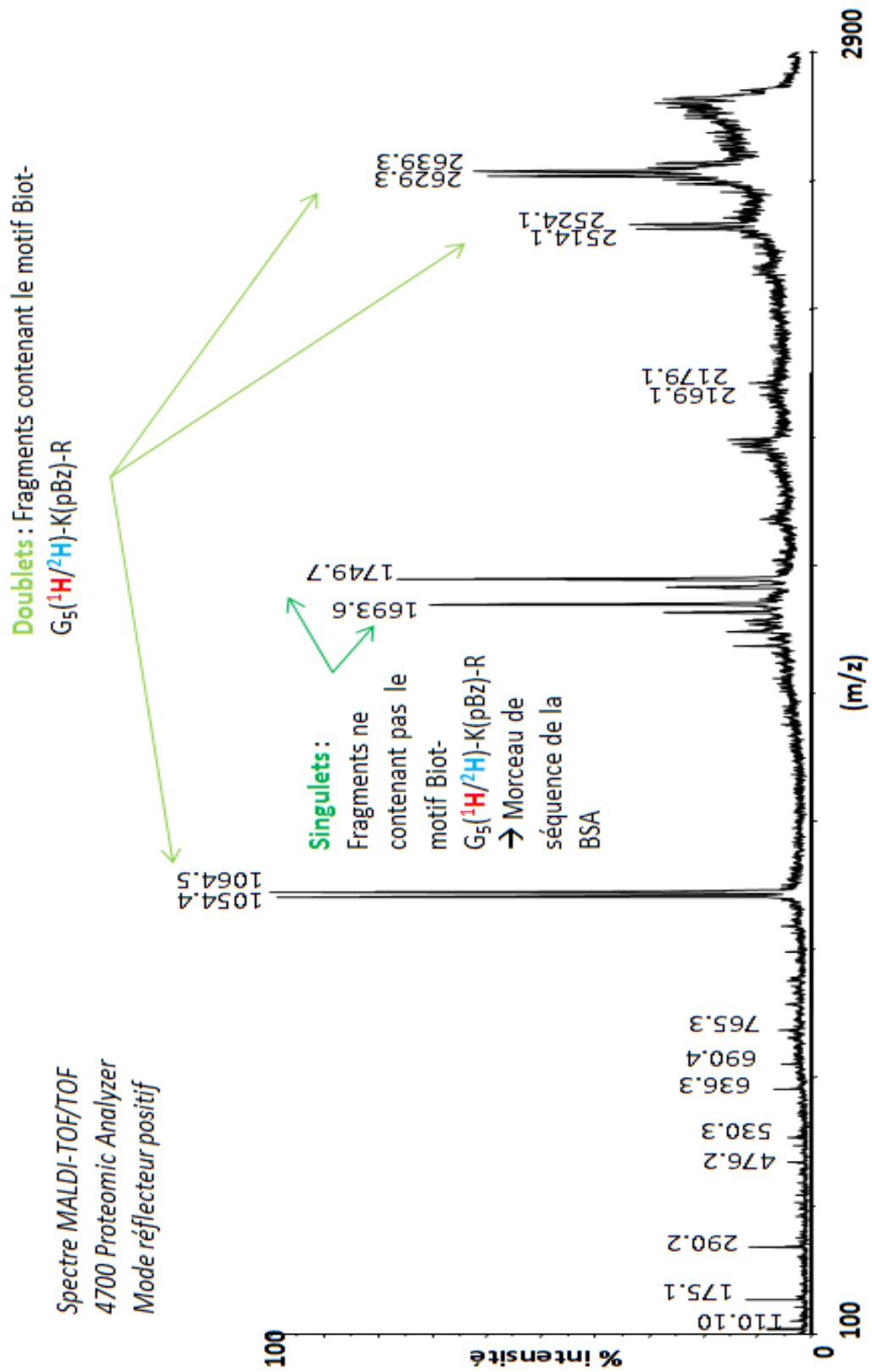


Spectres ETD (Zoom sur la gamme de m/z 150-1000) (ESI-LTQ/Orbitrap) des précurseurs 4<sup>+</sup> (701,78 et 704,29) pour les ions de (M+H)<sup>+</sup> 2803 et 2813.





Spectres ETD (Zoom sur la gamme de m/z 950-1500) (ESI-LTQ/Orbitrap) des précurseurs 4<sup>+</sup> (701,78 et 704,29) pour les ions de (M+H)<sup>+</sup> 2803 et 2813.



Spectre CID haute énergie (MALDI-TOF/TOF) des formes monochargées des ions de  $(M+H)^+$  2803 et 2813.

## Chapitre 6 : Pull down et photocross-linking

### Annexe II 3 : COMPARAISON A : Pull-down + photocross-linking versus contrôle négatif

N°	Bait	Prey	PreyGene	IP	Spec	SpecSum	NumRep	iProb	ctrlCounts	AvgP	MaxP
1	Pull-down +photoXlink	Q61782	keratin 14	Photo2 Photo1	13 15	28	2	1.00 1.00	0 0	1	1
2	Pull-down +photoXlink	P62908	ribosomal protein S3	Photo2 Photo1	33 49	82	2	0.98 1.00	10 0	0,99	1
3	Pull-down +photoXlink	O54724	polymerase I and transcript release factor	Photo2 Photo1	10 22	32	2	1.00 1.00	0 0	1	1
4	Pull-down +photoXlink	D3Z5M2	Protein Gm10110 RNA binding	Photo2 Photo1	8 14	22	2	1.00 1.00	0 0	1	1
5	Pull-down +photoXlink	P09405	nucleolin	Photo2 Photo1	49 78	127	2	1.00 1.00	12 0	0,9987	1
6	Pull-down +photoXlink	P29341	poly(A) binding protein, cytoplasmic 1, polyadenylate binding protein mRNA metabolism	Photo2 Photo1	9 18	27	2	1.00 1.00	0 0	1	1
7	Pull-down +photoXlink	Q9D8E6	ribosomal protein L4	Photo2 Photo1	20 34	54	2	1.00 1.00	0 0	1	1
8	Pull-down +photoXlink	P14148	60S ribosomal protein L7	Photo2 Photo1	19 34	53	2	0.97 1.00	4 1	0,9862	1
9	Pull-down +photoXlink	Q922F4	tubulin, beta 6	Photo2 Photo1	57 53	110	2	1.00 1.00	0 0	1	1
10	Pull-down +photoXlink	Q8R5L1	complement component 1, q subcomponent binding protein	Photo2 Photo1	49 96	145	2	1.00 1.00	0 0	1	1
11	Pull-down +photoXlink	P63101	tyrosine 3- monooxygenase/tryptop han 5-monooxygenase activation protein, zeta polypeptide;	Photo2 Photo1	8 8	16	2	1.00 1.00	0 0	0,9987	1
12	Pull-down +photoXlink	D3YVE6	unknown	Photo2 Photo1	22 33	55	2	0.98 1.00	5 2	0,9887	0,9975
13	Pull-down +photoXlink	P62259	tyrosine 3- monooxygenase/tryptop han 5-monooxygenase activation protein, epsilon polypeptide	Photo2 Photo1	14 8	22	2	1.00 1.00	0 0	0,9975	0,9975

Comparaison A : Protéines partenaires potentielles identifiées pour la condition Pull-down + Photocrosslinking (1<sup>ère</sup> partie de la liste/4)

N°	Bait	Prey	PreyGene	IP	Spec	SpecSum	NumRep	iProb	ctrlCounts	AvgP	MaxP
14	Pull-down +photoXlink	p97855	Ras-GTPase-activating protein SH3-domain binding protein 1	Photo2 Photo1	3 8	11	2	0.94 1.00	0 0	0,9688	0,995
15	Pull-down +photoXlink	p68040	guanine nucleotide binding protein (G protein), beta polypeptide 2 like 1	Photo2 Photo1	5 8	13	2	0.99 1.00	0 0	0,9925	0,995
16	Pull-down +photoXlink	B0V2N7	Annexin A2	Photo2 Photo1	27 18	45	2	0.99 0.96	6 0	0,975	0,99
17	Pull-down +photoXlink	Q9D0I9	arginyl-tRNA synthetase	Photo2 Photo1	7 5	12	2	0.99 0.99	0 0	0,99	0,99
18	Pull-down +photoXlink	Q545F8	ribosomal protein S4, X- linked	Photo2 Photo1	5 7	12	2	0.98 0.99	0 0	0,9875	0,99
19	Pull-down +photoXlink	O08583	THO complex 4	Photo2 Photo1	3 7	10	2	0.95 0.98	0 0	0,97	0,985
20	Pull-down +photoXlink	Q9D880	translocase of inner mitochondrial membrane 50 homolog (yeast)	Photo2 Photo1	4 3	7	2	0.98 0.96	0 0	0,9725	0,98
21	Pull-down +photoXlink	Q3TG12	phenylalanyl-tRNA synthetase, beta subunit	Photo2 Photo1	3 4	7	2	0.96 0.98	0 0	0,97	0,98
22	Pull-down +photoXlink	Q61033	thymopoietin	Photo2 Photo1	8 5	13	2	0.98 0.97	0 0	0,9762	0,98
23	Pull-down +photoXlink	p31230	aminoacyl-tRNA synthetase complex- interacting multifunctional protein 1	Photo2 Photo1	5 5	10	2	0.97 0.97	0 0	0,9737	0,975
24	Pull-down +photoXlink	Q5RKP3	ribosomal protein L13	Photo2 Photo1	7 12	19	2	0.89 0.97	2 0	0,9337	0,9725
25	Pull-down +photoXlink	Q9Z237	60S ribosomal protein L8	Photo2 Photo1	2 5	7	2	0.85 0.97	0 0	0,9112	0,97
26	Pull-down +photoXlink	P15864	histone cluster 1, H1c	Photo2 Photo1	8 6	14	2	0.96 0.95	1 0	0,9538	0,9575
27	Pull-down +photoXlink	B7ZNW0	Pgam5 protein Phosphoglycerate mutase family member 5, isoform CRA_b	Photo2 Photo1	3 2	5	2	0.95 0.87	0 0	0,91	0,955

Comparaison A : Protéines partenaires potentielles identifiées pour la condition Pull-down + Photocrosslinking (2<sup>ème</sup> partie de la liste/4).

N°	Bait	Prey	PreyGene	IP	Spec	SpecSum	NumRep	iProb	ctrlCounts	AvgP	MaxP
28	Pull-down +photoXlink	Q3THU8	solute carrier family 25 (mitochondrial carrier, phosphate carrier), member 3	Photo2 Photo1	2 4	6	2	0.87 0.95	0 0	0,9112	0,955
29	Pull-down +photoXlink	P11983	t-complex protein 1	Photo2 Photo1	3 3	6	2	0.95 0.93	0 0	0,9412	0,9525
30	Pull-down +photoXlink	P62830	ribosomal protein L23	Photo2 Photo1	7 5	12	2	0.95 0.93	1 0	0,9412	0,95
31	Pull-down +photoXlink	O88477	insulin-like growth factor 2 mRNA binding protein 1	Photo2 Photo1	2 4	6	2	0.88 0.94	0 0	0,9125	0,94
32	Pull-down +photoXlink	P43274	histone cluster1, H1e	Photo2 Photo1	3 3	6	2	0.94 0.94	0 0	0,9362	0,9375
33	Pull-down +photoXlink	Q5RKP4	ribophorin I	Photo2 Photo1	3 2	5	2	0.92 0.89	0 0	0,9087	0,9225
34	Pull-down +photoXlink	P51881	ADP/ATP translocase 2	Photo2 Photo1	43 26	69	2	0.96 0.84	13 1	0,9013	0,96
35	Pull-down +photoXlink	P27659	60S ribosomal protein L3	Photo2 Photo1	3 13	16	2	0.82 0.98	0 0	0,9	0,9825
36	Pull-down +photoXlink	G3UZ44	Paired mesoderm homeobox protein 1	Photo2 Photo1	2 2	4	2	0.89 0.90	0 0	0,8987	0,905
37	Pull-down +photoXlink	Q8BMS1	Trifunctional enzyme subunit alpha, mitochondrial	Photo2 Photo1	6 11	17	2	0.86 0.94	2 0	0,8975	0,935
38	Pull-down +photoXlink	B9EIU1	Glutamyl-prolyl-tRNA synthetase	Photo2 Photo1	2 3	5	2	0.86 0.93	0 0	0,895	0,9275
39	Pull-down +photoXlink	Q99L45	Eukaryotic translation initiation factor 2 subunit 2	Photo2 Photo1	2 2	4	2	0.89 0.90	0 0	0,8925	0,8975
40	Pull-down +photoXlink	Q8R010	Aminoacyl tRNA synthase complex-interacting multifunctional protein 2	Photo2 Photo1	2 2	4	2	0.87 0.87	0 0	0,8663	0,8675
41	Pull-down +photoXlink	Q3UEI6	Putative uncharacterized protein Serbp1	Photo2 Photo1	3 9	12	2	0.74 0.94	1 0	0,8438	0,945
42	Pull-down +photoXlink	O35129	Prohibitin-2	Photo2 Photo1	16 13	29	2	0.85 0.81	7 0	0,83	0,8525
43	Pull-down +photoXlink	Q61833	Ribophorin	Photo2 Photo1	4 3	7	2	0.85 0.80	1 0	0,8262	0,85

Comparaison A : Protéines partenaires potentielles identifiées pour la condition Pull-down + Photocrosslinking (3<sup>ème</sup> partie de la liste/4).

N°	Bait	Prey	PreyGene	IP	Spec	SpecSum	NumRep	iProb	ctrlCounts	AvgP	MaxP
44	Pull-down +photoXlink	Q9Z2K1	Keratin, type I cytoskeletal 16	Photo2 Photo1	12 18	30	2	0.76 0.87	7 0	0,8175	0,8725
45	Pull-down +photoXlink	P48962	ADP/ATP translocase 1	Photo2 Photo1	48 31	79	2	0.90 0.69	16 3	0,8	0,905
46	Pull-down +photoXlink	P70698	CTP synthase 1	Photo2 Photo1	3 4	7	2	0.77 0.82	1 0	0,7975	0,825
47	Pull-down +photoXlink	P16858	Glyceraldehyde-3-phosphate dehydrogenas	Photo2 Photo1	81 90	171	2	0.75 0.77	52 0	0,7613	0,7675
48	Pull-down +photoXlink	B3RH24	Lamin C	Photo2 Photo1	7 2	9	2	0.88 0.62	1 0	0,7537	0,8825

Comparaison A : Protéines partenaires potentielles identifiées pour la condition Pull-down + Photocrosslinking (4<sup>ème</sup> partie de la liste/4).

**Annexe II 4 : COMPARAISON B : Pull-down versus contrôle négatif**

N°	Bait	Prey	PreyGene	IP	Spec	SpecSum	NumRep	iProb	ctrlCounts	AvgP	MaxP
1	Pull-down	A1L0X5	keratin 78	Co2 Co1	45 20	65	2	1.00 1.00	0 0	1	1
2	Pull-down	A2A513	keratin 10	Co2 Co1	91 71	162	2	1.00 1.00	0 0	1	1
3	Pull-down	D3YVE6	Uncharacterized protein	Co2 Co1	9 13	22	2	1.00 1.00	0 0	1	1
4	Pull-down	E9Q5F6	Polyubiquitin-C	Co2 Co1	9 7	16	2	1.00 1.00	0 0	1	1
5	Pull-down	P04104	keratin 1	Co2 Co1	136 93	229	2	1.00 1.00	0 0	1	1
6	Pull-down	P10126	eukaryotic translation elongation factor 1 alpha 1	Co2 Co1	29 21	50	2	1.00 1.00	0 0	1	1
7	Pull-down	P11499	heat shock protein 90 alpha (cytosolic), class B member 1	Co2 Co1	23 32	55	2	1.00 1.00	0 0	1	1
8	Pull-down	P14148	ribosomal protein L7	Co2 Co1	9 23	32	2	1.00 1.00	0 0	1	1
9	Pull-down	P16381	DNA segment, Chr 1, Pasteur Institute 1	Co2 Co1	9 41	50	2	1.00 1.00	0 0	1	1
10	Pull-down	P48962	solute carrier family 25 member 4	Co2 Co1	18 6	24	2	1.00 1.00	0 0	1	1
11	Pull-down	P51881	similar to ADP/ATP translocase 2	Co2 Co1	15 7	22	2	1.00 1.00	0 0	1	1
12	Pull-down	P68372	tubulin, beta 2c, pseudogene 1; tubulin, beta 2C; tubulin, beta 2c, pseudogene 2	Co2 Co1	32 34	66	2	1.00 1.00	0 0	1	1
13	Pull-down	P68373	tubulin, alpha 1C; predicted gene 6682	Co2 Co1	30 41	71	2	1.00 1.00	0 0	1	1
14	Pull-down	P99024	tubulin, beta 5	Co2 Co1	42 43	85	2	1.00 1.00	0 0	1	1
15	Pull-down	Q03265	ATP synthase, H+ transporting, mitochondrial F1 complex, alpha subunit, isoform 1	Co2 Co1	12 9	21	2	1.00 1.00	0 0	1	1
16	Pull-down	Q3TUZ9	glyceraldehyde-3-phosphate dehydrogenase	Co2 Co1	29 19	48	2	1.00 1.00	0 0	1	1
17	Pull-down	Q61033	thymopoietin	Co2 Co1	8 17	25	2	1.00 1.00	0 0	1	1
18	Pull-down	Q6IFX2	keratin 42	Co2 Co1	76 31	107	2	1.00 1.00	0 0	1	1
19	Pull-down	Q6NXH9	keratin 73	Co2 Co1	102 52	154	2	1.00 1.00	0 0	1	1

Comparaison B : Protéines partenaires potentielles identifiées pour la condition Pull-down (1<sup>ère</sup> partie de la liste/3).



N°	Bait	Prey	PreyGene	IP	Spec	SpecSum	NumRep	iProb	ctrlCounts	AvgP	MaxP
20	Pull-down	Q71V06	Y box protein 1	Co2 Co1	18 27	45	2	1.00 1.00	0 0	1	1
21	Pull-down	Q8R5L1	complement component 1, q subcomponent binding protein	Co2 Co1	30 63	93	2	1.00 1.00	0 0	1	1
22	Pull-down	Q8VED5	keratin 79	Co2 Co1	59 33	92	2	1.00 1.00	0 0	1	1
23	Pull-down	Q9CR57	ribosomal protein L14	Co2 Co1	10 9	19	2	1.00 1.00	0 0	1	1
24	Pull-down	Q9CSM5	ribosomal protein L4	Co2 Co1	7 19	26	2	1.00 1.00	0 0	1	1
25	Pull-down	Q9CV72	keratin 68	Co2 Co1	100 46	146	2	1.00 1.00	0 0	1	1
26	Pull-down	Q9Z1R9	protease, serine, 1 (trypsin 1)	Co2 Co1	60 50	110	2	1.00 1.00	0 0	1	1
27	Pull-down	E9Q3T0	Uncharacterized protein Gm10073	Co2 Co1	7 7	14	2	1.00 1.00	0 0	1	1
28	Pull-down	B3RH24	lamin A	Co2 Co1	4 5	9	2	1.00 1.00	0 0	1	1
29	Pull-down	I7HLV2	ribosomal protein L10	Co2 Co1	8 5	13	2	1.00 1.00	0 0	0,9987	1
30	Pull-down	Q3TB63	heat shock protein 8	Co2 Co1	5 6	11	2	1.00 1.00	0 0	0,9975	1
31	Pull-down	Q5RKP3	similar to ribosomal protein L13	Co2 Co1	3 4	7	2	1.00 1.00	0 0	0,9975	1
32	Pull-down	Q8BMK4	cytoskeleton-associated protein 4	Co2 Co1	5 25	30	2	0.99 1.00	0 0	0,995	0,9975
33	Pull-down	P56480	ATP synthase, H+ transporting mitochondrial F1 complex, beta subunit	Co2 Co1	23 5	28	2	1.00 0.98	0 0	0,9925	1
34	Pull-down	Q5SUT0	Ewing sarcoma breakpoint region 1	Co2 Co1	4 4	8	2	0.99 0.99	0 0	0,99	0,9925
35	Pull-down	Q3V1Z7	DEAD (Asp-Glu-Ala-Asp) box polypeptide 1	Co2 Co1	2 5	7	2	0.95 0.99	0 0	0,9737	0,9925
36	Pull-down	P27659	similar to 60S ribosomal protein L3 (J1 protein)	Co2 Co1	2 8	10	2	0.94 1.00	0 0	0,9688	0,995
37	Pull-down	D3Z6C3	40S ribosomal protein S3a	Co2 Co1	6 2	8	2	0.99 0.95	0 0	0,9688	0,9875
38	Pull-down	P62908	ribosomal protein S3	Co2 Co1	3 14	17	2	0.94 1.00	0 0	0,9675	0,9975
39	Pull-down	Q3UD36	vimentin	Co2 Co1	2 3	5	2	0.95 0.98	0 0	0,965	0,9825

Comparaison B : Protéines partenaires potentielles identifiées pour la condition Pull-down (2<sup>ème</sup> partie de la liste/3).

N°	Bait	Prey	PreyGene	IP	Spec	SpecSum	NumRep	iProb	ctrlCounts	AvgP	MaxP
40	Pull-down	Q61753	3-phosphoglycerate dehydrogenase	Co2 Co1	3 14	17	2	0.92 1.00	0 0	0,96	0,9975
41	Pull-down	Q9CV24	similar to Eukaryotic translation initiation factor 2 subunit 1	Co2 Co1	2 2	4	2	0.96 0.94	0 0	0,9512	0,9575
42	Pull-down	Q5RL55	ATP-binding cassette, sub-family F (GCN20), member 1	Co2 Co1	2 10	12	2	0.89 1.00	0 0	0,945	0,9975
43	Pull-down	B7STB7	Annexin 1	Co2 Co1	8 2	10	2	0.98 0.88	0 0	0,9287	0,98
44	Pull-down	P07724	albumin	Co2 Co1	2 14	16	2	0.68 1.00	0 0	0,8375	0,995
45	Pull-down	Q3UDA7	nucleophosmin 1	Co2 Co1	1 6	7	2	0.69 0.96	0 0	0,8225	0,96
46	Pull-down	Q89072	ribosomal protein S2	Co2 Co1	1 6	7	2	0.66 0.98	0 0	0,82	0,9775
47	Pull-down	E9QN31	Putative ribosomal RNA methyltransferase NOP2	Co2 Co1	1 2	3	2	0.71 0.93	0 0	0,8175	0,9275
48	Pull-down	P62754	similar to 40S ribosomal protein S6	Co2 Co1	1 2	3	2	0.72 0.91	0 0	0,8175	0,9125
49	Pull-down	E9Q132	60S ribosomal protein L24	Co2 Co1	1 5	6	2	0.68 0.95	0 0	0,8138	0,9525
50	Pull-down	D3YVN7	Elongation factor Tu	Co2 Co1	1 2	3	2	0.71 0.91	0 0	0,8112	0,91
51	Pull-down	Q8BX68	high density lipoprotein (HDL) binding protein	Co2 Co1	1 2	3	2	0.73 0.89	0 0	0,81	0,89
52	Pull-down	Q9CZK0	sorting nexin 9	Co2 Co1	1 7	8	2	0.61 0.99	0 0	0,7975	0,9875
53	Pull-down	P63038	heat shock protein 1 (chaperonin)	Co2 Co1	1 2	3	2	0.72 0.88	0 0	0,7975	0,8775
54	Pull-down	P09405	nucleolin	Co2 Co1	12 107	119	2	0.60 0.99	0 0	0,7962	0,9925
55	Pull-down	D3Z7N2	Elongation factor 1-delta	Co2 Co1	2 1	3	2	0.90 0.68	0 0	0,79	0,8975
56	Pull-down	G3UW34	60S ribosomal protein L9	Co2 Co1	8 1	9	2	0.98 0.53	0 0	0,7587	0,9825

Comparaison B : Protéines partenaires potentielles identifiées pour la condition Pull-down (3ème partie de la liste/3).

**Annexe II 5:** Attributions des spectres MS/MS aux peptides [269-304] et [373-386] de la protéine.

## ➤ Peptide QDHPSSMGVYGQESGGFSGGENR

Peptide Summary								
Sequence: QDHPSSMGVYGQESGGFSGGENR, M7-Oxidation (15.99492 Da)								
Charge: +3, Monoisotopic m/z: 832.68604 Da (-1.54 mmu/-1.85 ppm), MH+: 2496.04355 Da, RT: 64.35 min,								
Identified with: SEQUEST (v1.20); XCorr:4.37, Probability:0.00, Ions matched by search engine: 28/92								
Fragment match tolerance used for search: 0.8 Da								
Fragment Matches								
Value Type: Theo. Mass [Da]								
Ion Series								
Neutral Losses								
Precursor Ions								
#1	b <sup>+</sup>	b <sup>2+</sup>	b <sup>3+</sup>	Seq.	y <sup>+</sup>	y <sup>2+</sup>	y <sup>3+</sup>	#2
1	129.06586	65.03657	43.69347	Q				24
2	244.09281	122.55004	82.03579	D	2367.98959	1184.49843	790.00138	23
3	381.15172	191.07950	127.72209	H	2252.96264	1126.98496	751.65906	22
4	478.20449	239.60588	160.07301	P	2115.90373	1058.45550	705.97276	21
5	565.23652	283.12190	189.08369	S	2018.85096	1009.92912	673.62184	20
6	652.26855	326.63791	218.09437	S	1931.81893	966.41310	644.61116	19
7	799.30396	400.15562	267.10617	M-Oxidation	1844.78690	922.89709	615.60048	18
8	856.32543	428.66635	286.11333	G	1697.75148	849.37938	566.58868	17
9	955.39385	478.20056	319.13613	V	1640.73001	820.86864	547.58152	16
10	1118.45717	559.73222	373.49057	Y	1541.66159	771.33443	514.55871	15
11	1175.47864	588.24296	392.49773	G	1378.59827	689.80277	460.20427	14
12	1303.53722	652.27225	435.18392	Q	1321.57680	661.29204	441.19712	13
13	1432.57982	716.79355	478.19812	E	1193.51822	597.26275	398.51092	12
14	1519.61185	760.30956	507.20880	S	1064.47562	532.74145	355.49672	11
15	1576.63332	788.82030	526.21596	G	977.44359	489.22543	326.48605	10
16	1633.65479	817.33103	545.22311	G	920.42212	460.71470	307.47889	9
17	1780.72321	890.86524	594.24592	F	863.40065	432.20396	288.47173	8
18	1867.75524	934.38126	623.25660	S	716.33223	358.66975	239.44893	7

## ➤ Peptide GDATVSYEDP

Peptide Summary						
Sequence: GDATVSYEDPPTAK, Charge: +2, Monoisotopic m/z: 725.83588 Da (-2.61 mmu/-3.59 ppm),						
Identified with: SEQUEST (v1.20); XCorr:2.64, Probability:0.00, Ions matched by search engine: 15/26						
Fragment match tolerance used for search: 0.8 Da						
Fragments used for search: b; b-H <sub>2</sub> O; y; y-H <sub>2</sub> O; y-NH <sub>3</sub>						
Fragment Matches						
Value Type: Theo. Mass [Da]						
Ion Series						
Neutral Losses						
Precursor Ions						
#1	b <sup>+</sup>	b <sup>2+</sup>	Seq.	y <sup>+</sup>	y <sup>2+</sup>	#2
1	58.02875	29.51801	G			14
2	173.05570	87.03149	D	1393.64822	697.32775	13
3	244.09282	122.55005	A	1278.62127	639.81427	12
4	345.14050	173.07389	T	1207.58415	604.29571	11
5	444.20892	222.60810	V	1106.53647	553.77187	10
6	531.24095	266.12411	S	1007.46805	504.23766	9
7	694.30427	347.65577	Y	920.43602	460.72165	8
8	823.34687	412.17707	E	757.37270	379.18999	7
9	938.37382	469.69055	D	628.33010	314.66869	6
10	1035.42659	518.21693	P	513.30315	257.15521	5
11	1132.47936	566.74332	P	416.25038	208.62883	4
12	1233.52704	617.26716	T	319.19761	160.10244	3
13	1304.56416	652.78572	A	218.14993	109.57860	2
14			K	147.11281	74.06004	1



## Chapitre 7 : Protéomique quantitative différentielle

**Annexe II 6:** Liste de protéines sur- ou sous-exprimées dans les cellules EF (H) par rapport aux cellules 3T3

50/50 proteins	EF-3T3-AL...avy/Light				EFvs3T3-2...avy/Light				EF-3T3_al...avy/Light				MW kDa	Description - Species
	Ratio	Pep. used	Ratio	Pep. used	Ratio	Pep. used	Ratio	Pep. used	Ratio	Pep. used	Ratio	Pep. used		
ADT1_MOUSE	-	-	1.83 ⬆	9	1.83 ⬆	33	32.9	ADP/ATP translocase 1 <i>Mus musculus</i>						
ADT2_MOUSE	1.75	2	1.83 ⬆	9	1.79 ⬆	37	32.9	ADP/ATP translocase 2 <i>Mus musculus</i>						
ALDOA_MOUSE	1.79 ⬆	8	1.50 ⬆	29	1.78 ⬆	38	39.4	Fructose-bisphosphate aldolase A <i>Mus musculus</i>						
ANX11_MOUSE	-	-	2.41 ⬆	4	3.38 ⬆	2	54.1	Annexin A11 <i>Mus musculus</i>						
ANXA3_MOUSE	3.73 ⬆	6	4.48 ⬆	15	4.59 ⬆	26	36.4	Annexin A3 <i>Mus musculus</i>						
C11C_MOUSE	2.08 ⬆	2	2.06 ⬆	15	1.51 ⬆	9	101.2	C-1-tetrahydrofolate synthase, cytoplasmic <i>Mus musculus</i>						
CBR3_MOUSE	1.47 ⬆	2	1.73 ⬆	4	1.69 ⬆	4	31.0	Carbonyl reductase [NADPH] 3 <i>Mus musculus</i>						
CLIC1_MOUSE	1.62 ⬆	4	1.76 ⬆	16	2.02 ⬆	24	27.0	Chloride intracellular channel protein 1 <i>Mus musculus</i>						
D3Z3R1_MOUSE	1.57 ⬆	3	1.50	5	1.23	2	12.1	60S ribosomal protein L36 <i>Mus musculus</i>						
DYR_MOUSE	-	-	1.73 ⬆	5	1.57	1	21.6	Dihydrofolate reductase <i>Mus musculus</i>						
E9QN31_MOUSE	10.01 ⬆	1	2.28 ⬆	5	-	-	86.9	Puative ribosomal RNA methyltransferase NOP2 <i>Mus musculus</i>						
EIF3C_MOUSE	1.41	1	1.61 ⬆	3	1.73	4	105.5	Eukaryotic translation initiation factor 3 subunit C <i>Mus musculus</i>						
EIF3L_MOUSE	-	-	1.66 ⬆	9	1.67	3	66.6	Eukaryotic translation initiation factor 3 subunit L <i>Mus musculus</i>						
FAS_MOUSE	1.73 ⬆	19	1.60 ⬆	22	1.55 ⬆	35	272.4	Fatty acid synthase <i>Mus musculus</i>						
G3P_MOUSE	1.70 ⬆	22	1.46 ⬆	104	1.57 ⬆	111	35.8	Glyceraldehyde-3-phosphate dehydrogenase <i>Mus musculus</i>						
G6PD1_MOUSE	-	-	2.93 ⬆	2	3.14 ⬆	8	59.3	Glucose-6-phosphate 1-dehydrogenase X <i>Mus musculus</i>						
G6P1_MOUSE	2.13 ⬆	4	1.85 ⬆	9	1.55 ⬆	15	62.8	Glucose-6-phosphate isomerase <i>Mus musculus</i>						
GRPE1_MOUSE	-	-	2.35 ⬆	3	2.53 ⬆	3	24.3	GrpE protein homolog 1, mitochondrial <i>Mus musculus</i>						
GSTO1_MOUSE	∞ ⬆	1	-	-	2.21 ⬆	8	27.5	Glutathione S-transferase omega-1 <i>Mus musculus</i>						
HYOU1_MOUSE	1.53	1	1.23 ⬆	12	1.61 ⬆	16	111.2	Hypoxia up-regulated protein 1 <i>Mus musculus</i>						
IF2B1_MOUSE	-	-	1.83 ⬆	6	1.63	2	63.5	Insulin-like growth factor 2 mRNA-binding protein 1 <i>Mus musculus</i>						
IF2G_MOUSE	-	-	1.58 ⬆	4	∞ ⬆	1	51.1	Eukaryotic translation initiation factor 2 subunit 3, X-linked <i>Mus musculus</i>						
ITPA1_MOUSE	2.93 ⬆	3	2.44 ⬆	10	2.95 ⬆	3	57.9	Importin subunit alpha-1 <i>Mus musculus</i>						
KPYM_MOUSE	1.56	4	1.44 ⬆	126	1.58 ⬆	203	57.8	Pyruvate kinase PKM <i>Mus musculus</i>						
LDHA_MOUSE	1.69 ⬆	2	2.01 ⬆	14	1.50 ⬆	20	36.5	L-lactate dehydrogenase A chain <i>Mus musculus</i>						

Protéines surexprimées dans les cellules EF (H) par rapport aux cellules 3T3 (protéines avec une variation du ratio H/L supérieure à 1,5 et une p-value inférieure à 0,05 pour au moins deux des trois réplicats). 1<sup>ère</sup> partie/2 de la liste

50/50 proteins	EF-3T3-AL...avy/Light EFvs3T3-2...avy/Light EF-3T3_al...avy/Light						MW kDa	Description - Species
	Ratio	Pep. used	Ratio	Pep. used	Ratio	Pep. used		
MBB1A_MOUSE	1.48 ↑	13	1.66 ↑	29	1.35 ↑	39	152.0	Myb-binding protein 1A <u>Mus musculus</u>
MCM6_MOUSE	2.28 ↑	1	2.44 ↑	1	2.46 ↑	5	92.9	DNA replication licensing factor MCM6 <u>Mus musculus</u>
MOES_MOUSE	-	-	1.42 ↑	8	1.60 ↑	18	67.8	Moiesin <u>Mus musculus</u>
NASP_MOUSE	9.14 ↑	1	6.49 ↑	5	-	-	84.0	Nuclear autoantigenic sperm protein <u>Mus musculus</u>
NCPR_MOUSE	-	-	∞ ↑	5	5.19 ↑	7	77.0	NADPH--cytochrome P450 reductase <u>Mus musculus</u>
PAIRB_MOUSE	1.79	1	2.22 ↑	6	1.67	2	44.7	Plasminogen activator inhibitor 1 RNA-binding protein <u>Mus musculus</u>
PCNA_MOUSE	-	-	1.97 ↑	10	2.10 ↑	25	28.8	Proliferating cell nuclear antigen <u>Mus musculus</u>
PFD2_MOUSE	1.80 ↑	3	1.54 ↑	4	1.37	3	16.5	Prefoldin subunit 2 <u>Mus musculus</u>
PGAM1_MOUSE	1.69 ↑	6	1.47 ↑	25	1.43 ↑	31	28.8	Phosphoglycerate mutase 1 <u>Mus musculus</u>
PGK1_MOUSE	1.39 ↑	17	1.42 ↑	34	1.62 ↑	63	44.6	Phosphoglycerate kinase 1 <u>Mus musculus</u>
PURA2_MOUSE	2.44 ↑	1	2.95 ↑	5	2.49 ↑	2	50.0	Adenylosuccinate synthetase isozyme 2 <u>Mus musculus</u>
Q3UFZ6_MOUSE	1.29 ↑	6	1.57 ↑	12	1.76	4	37.6	Putative uncharacterized protein (Fragment) <u>Mus musculus</u>
Q5U415_MOUSE	4.36 ↑	2	3.17 ↑	5	4.71 ↑	6	35.7	Aldo-keto reductase family 1, member B3 (Aldose reductase) <u>Mus musculus</u>
Q91YT9_MOUSE	1.53 ↑	2	1.81 ↑	7	-	-	33.7	6-phosphogluconate dehydrogenase, decarboxylating <u>Mus musculus</u>
RL3_MOUSE	1.42 ↑	2	2.15 ↑	2	1.75 ↑	10	46.1	60S ribosomal protein L3 <u>Mus musculus</u>
RL7A_MOUSE	-	-	1.52 ↑	12	1.67 ↑	10	30.0	60S ribosomal protein L7a <u>Mus musculus</u>
SAHH_MOUSE	1.69 ↑	6	1.90 ↑	11	1.52 ↑	22	47.7	Adenosylhomocysteinase <u>Mus musculus</u>
SLK_MOUSE	1.84 ↑	3	3.81 ↑	3	∞ ↑	3	141.5	STE20-like serine/threonine-protein kinase <u>Mus musculus</u>
SPB6_MOUSE	-	-	1.54 ↑	5	1.62 ↑	16	42.6	Serpin B6 <u>Mus musculus</u>
SSRD_MOUSE	-	-	1.50	4	2.12 ↑	7	18.9	Translocon-associated protein subunit delta <u>Mus musculus</u>
SYCC_MOUSE	1.56	1	1.55	2	2.47 ↑	4	94.9	Cysteine--tRNA ligase, cytoplasmic <u>Mus musculus</u>
SVC_MOUSE	1.81 ↑	9	1.95 ↑	24	1.62 ↑	16	140.2	Valine--tRNA ligase <u>Mus musculus</u>
XPO2_MOUSE	-	-	1.91 ↑	4	1.42 ↑	9	110.5	Exportin-2 <u>Mus musculus</u>
YBOX1_MOUSE	1.56 ↑	5	1.47	6	1.34 ↑	14	35.7	Nuclease-sensitive element-binding protein 1 <u>Mus musculus</u>
YBOX3_MOUSE	8.21 ↑	2	2.19 ↑	4	1.83 ↑	12	38.8	Y-box-binding protein 3 <u>Mus musculus</u>

Protéines surexprimées dans les cellules EF (H) par rapport aux cellules 3T3 (protéines avec une variation du ratio H/L supérieure à 1,5 et une p-valeur inférieure à 0,05 pour au moins deux des trois réplicats). 1<sup>ère</sup> partie/2 de la liste

56/56 proteins	EF-3T3-Al...avy/Light				EFvs3T3-2...avy/Light				EF-3T3_al...avy/Light				MW kDa	Description - Species
	Ratio	Pep. used	Ratio	Pep. used	Ratio	Pep. used	Ratio	Pep. used	Ratio	Pep. used	Ratio	Pep. used		
1433T_MOUSE	1/1.73	2	1/1.56	13	1/1.44	21	27.8	14-3-3 protein theta <i>Mus musculus</i>						
A2A6G6_MOUSE	1/1.86	3	1/2.00	4	1/1.87	4	10.6	LIM and SH3 domain protein 1 (Fragment) <i>Mus musculus</i>						
ACAD9_MOUSE	-	-	1/1.95	4	1/2.48	6	68.7	Acyl-CoA dehydrogenase family member 9, mitochondrial <i>Mus musculus</i>						
AK1A1_MOUSE	1/1.70	1	1/1.64	4	1/1.41	11	36.6	Alcohol dehydrogenase [NADP(+)] <i>Mus musculus</i>						
AL1L2_MOUSE	1/1.77	2	1/1.79	8	1/2.52	3	101.6	Mitochondrial 10-formyltetrahydrofolate dehydrogenase <i>Mus musculus</i>						
B9EKP5_MOUSE	1/1.76	39	1/2.09	75	1/2.46	111	280.5	Filamin, alpha <i>Mus musculus</i>						
BASP1_MOUSE	1/2.74	4	1/2.38	5	1/1.95	5	22.1	Brain acid soluble protein 1 <i>Mus musculus</i>						
CNPY2_MOUSE	1/1.98	1	1/1.89	5	1/1.21	2	20.8	Protein canopy homolog 2 <i>Mus musculus</i>						
D3Z4A4_MOUSE	1/1.93	2	1/1.67	3	1/1.73	12	16.1	Peroxiredoxin-2 (Fragment) <i>Mus musculus</i>						
DEST_MOUSE	-	-	1/1.91	2	1/2.36	13	18.5	Destrin <i>Mus musculus</i>						
ETFA_MOUSE	1/1.55	1	1/1.35	10	1/1.55	17	35.0	Electron transfer flavoprotein subunit alpha, mitochondrial <i>Mus musculus</i>						
F7CVJ5_MOUSE	1/1.65	10	1/1.67	23	1/1.18	33	115.3	Protein Ahnak2 (Fragment) <i>Mus musculus</i>						
F7DBB3_MOUSE	1/1.70	8	1/1.66	22	1/1.15	35	166.5	Protein Ahnak2 (Fragment) <i>Mus musculus</i>						
FRIL1_MOUSE	1/4.65	4	-	-	1/1.71	22	20.8	Ferritin light chain 1 <i>Mus musculus</i>						
GELS_MOUSE	1/2.29	14	1/2.43	19	1/3.19	41	85.9	Gelsolin <i>Mus musculus</i>						
GLCM_MOUSE	-	-	1/2.49	3	1/1.89	2	57.6	Glucosylceramidase <i>Mus musculus</i>						
GNAI2_MOUSE	1/1.31	2	1/2.54	5	1/2.22	4	40.5	Guanine nucleotide-binding protein G(i) subunit alpha-2 <i>Mus musculus</i>						
GSTM1_MOUSE	-	-	1/∞	4	1/2.99	8	26.0	Glutathione S-transferase Mu 1 <i>Mus musculus</i>						
GSTM2_MOUSE	-	-	1/1.91	3	1/2.19	16	25.7	Glutathione S-transferase Mu 2 <i>Mus musculus</i>						
HMGAI1_MOUSE	1/1.63	5	1/1.40	14	1/1.77	9	11.6	High mobility group protein HMG-I/HMG-Y <i>Mus musculus</i>						
HMOX1_MOUSE	1/1.38	1	1/3.17	3	1/∞	1	32.9	Heme oxygenase 1 <i>Mus musculus</i>						
HXX1_MOUSE	1/1.60	2	-	-	1/2.17	6	108.3	Hexokinase-1 <i>Mus musculus</i>						
IDHC_MOUSE	1/2.44	1	1/1.85	2	1/1.70	11	46.7	Isocitrate dehydrogenase [NADP] cytoplasmic <i>Mus musculus</i>						
IDHP_MOUSE	1/3.75	1	1/1.90	9	1/1.73	8	50.9	Isocitrate dehydrogenase [NADP], mitochondrial <i>Mus musculus</i>						
IFM3_MOUSE	1/1.10	1	1/2.87	4	1/∞	5	15.0	Interferon-induced transmembrane protein 3 <i>Mus musculus</i>						

Protéines sous-exprimées dans les cellules EF (H) par rapport aux cellules 3T3 (protéines avec une variation du ratio H/L supérieure à 1,5 et une p-value inférieure à 0,05 pour au moins deux des trois réplicats). 1ère partie/2 de la liste

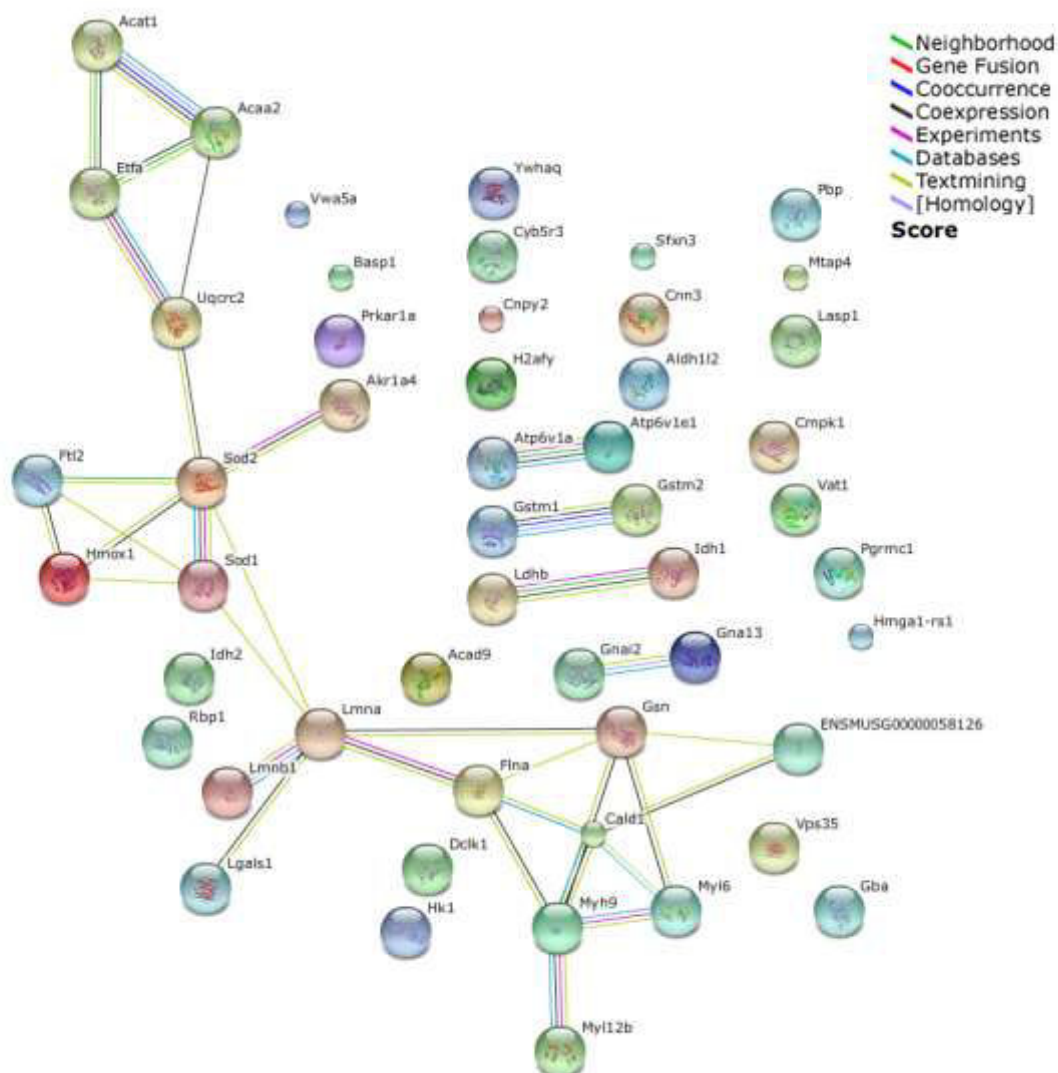


KAPO_MOUSE	-	1/1.88	3	1/1.51	2	43.2	cAMP-dependent protein kinase type I-alpha regulatory subunit <i>Mus musculus</i>	
KCY_MOUSE	1/2.45	2	1/2.03	2	1/2.04	4	22.2	UMP-CMP kinase <i>Mus musculus</i>
LDHB_MOUSE	1/6.32	4	1/3.76	9	1/2.89	20	36.6	L-lactate dehydrogenase B chain <i>Mus musculus</i>
LEGI_MOUSE	-	-	1/1.70	34	1/1.75	16	14.9	Galectin-1 <i>Mus musculus</i>
LMNA_MOUSE	1/2.33	27	1/2.58	104	1/2.42	110	74.2	Prelamin-A/C <i>Mus musculus</i>
LMNB1_MOUSE	-	-	1/1.46	6	1/1.55	7	66.8	Lamin-B1 <i>Mus musculus</i>
MAP4_MOUSE	1/1.11	7	1/1.67	5	1/1.69	11	117.4	Microtubule-associated protein 4 <i>Mus musculus</i>
ML12B_MOUSE	1/1.42	2	1/1.58	7	-	-	19.8	Myosin regulatory light chain 12B <i>Mus musculus</i>
MYH9_MOUSE	1/1.50	68	1/1.58	137	1/1.57	157	226.4	Myosin-9 <i>Mus musculus</i>
MYL6_MOUSE	1/1.42	4	1/1.53	17	1/1.44	13	16.9	Myosin light polypeptide 6 <i>Mus musculus</i>
NBSR3_MOUSE	-	-	1/2.36	18	1/2.19	30	34.1	NADH-cytochrome b5 reductase 3 <i>Mus musculus</i>
PEBP1_MOUSE	1/2.36	3	1/1.47	21	1/1.94	15	20.8	Phosphatidylethanolamine-binding protein 1 <i>Mus musculus</i>
PGR1_MOUSE	1/3.01	2	1/3.06	4	1/2.33	7	21.7	Membrane-associated progesterone receptor component 1 <i>Mus musculus</i>
Q3UGP0_MOUSE	1/5.91	3	1/7.40	10	1/∞	4	40.4	Putative uncharacterized protein <i>Mus musculus</i>
Q58E70_MOUSE	1/1.31	16	1/1.58	15	1/1.74	21	29.0	Tpm3 protein <i>Mus musculus</i>
Q8CA90_MOUSE	1/1.41	1	1/1.30	1	1/1.58	5	21.0	Histone H2A <i>Mus musculus</i>
Q8VCQ8_MOUSE	-	-	1/2.29	5	1/1.63	6	60.5	Caldesmon 1 <i>Mus musculus</i>
Q99JN5_MOUSE	1/1.84	2	-	-	1/2.52	4	20.1	Cnn3 protein (Fragment) <i>Mus musculus</i>
Q9D034_MOUSE	1/2.51	1	-	-	1/1.62	3	19.9	Guanine nucleotide-binding protein subunit alpha-13 <i>Mus musculus</i>
QCR2_MOUSE	1/2.02	2	1/1.36	7	1/1.85	5	48.2	Cytochrome b-c1 complex subunit 2, mitochondrial <i>Mus musculus</i>
RET1_MOUSE	1/23.21	1	1/5.44	6	1/3.21	3	15.8	Retinol-binding protein 1 <i>Mus musculus</i>
SFXN3_MOUSE	1/4.29	2	1/4.93	4	1/3.43	4	35.4	Sideroflexin-3 <i>Mus musculus</i>
SODC_MOUSE	1/2.00	6	1/1.53	6	1/1.43	17	15.9	Superoxide dismutase [Cu-Zn] <i>Mus musculus</i>
SODM_MOUSE	1/1.28	1	1/1.71	3	1/1.52	4	24.6	Superoxide dismutase [Mn], mitochondrial <i>Mus musculus</i>
THIL_MOUSE	1/1.27	3	1/1.61	6	1/2.00	12	44.8	Acetyl-CoA acetyltransferase, mitochondrial <i>Mus musculus</i>
THIM_MOUSE	-	-	1/2.95	7	1/∞	1	41.8	3-ketoacyl-CoA thiolase, mitochondrial <i>Mus musculus</i>
VAT1_MOUSE	1/2.90	5	1/3.10	24	1/1.49	29	43.1	Synaptic vesicle membrane protein VAT-1 homolog <i>Mus musculus</i>
VATA_MOUSE	-	-	1/1.95	10	1/1.70	7	68.3	V-type proton ATPase catalytic subunit A <i>Mus musculus</i>
VATE1_MOUSE	1/2.52	1	1/1.85	6	1/1.28	6	26.2	V-type proton ATPase subunit E 1 <i>Mus musculus</i>
VMA5A_MOUSE	1/3.10	1	1/2.45	4	1/4.65	2	87.1	von Willebrand factor A domain-containing protein 5A <i>Mus musculus</i>
VPS35_MOUSE	1/1.87	1	1/1.73	7	1/1.24	9	91.7	Vacuolar protein sorting-associated protein 35 <i>Mus musculus</i>

Protéines sous-exprimées dans les cellules EF (H) par rapport aux cellules 3T3 (protéines avec une variation du ratio H/L supérieure à 1,5 et une p-value inférieure à 0,05 pour au moins deux des trois réplicats). 2<sup>ème</sup> partie/2 de la liste.

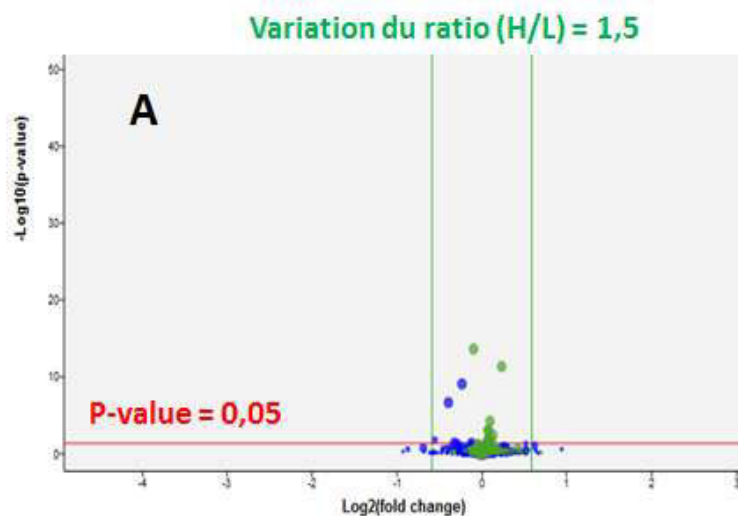


**Annexe II 7:** Réseaux d'interactions des protéines sous-exprimées

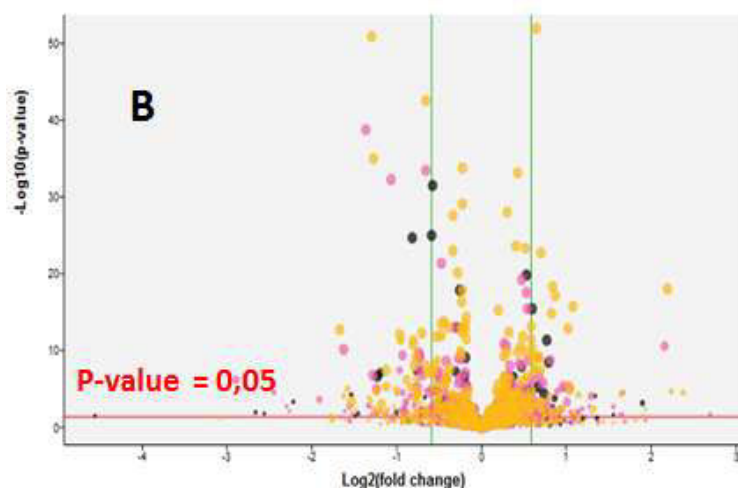


**Réseau d'interactions protéine-protéine généré sur le logiciel STRING à partir de la liste de protéines sous-exprimées dans les cellules EF par rapport aux cellules 3T3.** Chaque sphère correspond à une protéine et chaque trait correspond à un lien entre deux protéines dans la base de données. La couleur des traits fait référence au type de lien existant entre ces protéines comme indiqué dans la légende (en haut à droite)

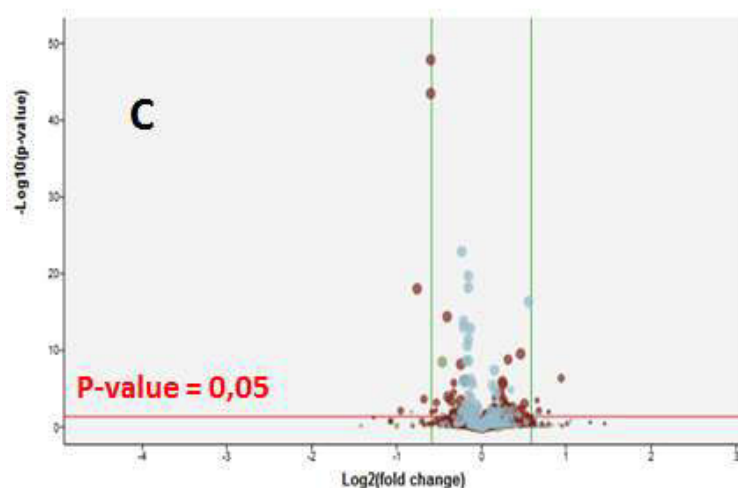
**Annexe II 8 :** Volcano plots obtenus pour les expériences de protéomique différentielle avec le marquage SILAC.



Volcano plot  
2 réplacats biologiques  
**contrôle = EF(H)/EF(L)**  
(1082 protéines identifiées  
et quantifiées)



Volcano plot  
3 réplacats biologiques  
**EF(H) / 3T3 (L)**  
(1661 protéines identifiées  
et quantifiées)



Volcano plot  
3 réplacats biologiques  
**EF(H) non traitées / EF(L)  
traitées avec 5  $\mu\text{M}$  de  
peptide photo(R/W)9**  
(1745 protéines identifiées  
et quantifiées)

Volcano plots obtenus pour les expériences de protéomique différentielle réalisées avec le marquage SILAC. La partie A correspond à l'expérience contrôle, la partie B à la comparaison EF(H)/3T3(L) et la partie C à la comparaison EF(H)/ EF(L) traitées par photo(R/W)<sub>9</sub>.

